

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6	10	1342.5	46.3	673	22	ABB66506
OM protein - protein search, using SW model.		11	1298.5	44.8	287	20	AAV60382
Run on:	January 28, 2004, 09:19:15 ; Search time 46 Seconds (without alignments)	12	936.5	32.3	593	21	AAAB43002
Title:	US-09-978-242-1	13	936.5	32.3	593	22	AAAM40333
Perfect score:	2898	14	895.5	30.9	524	22	AAAM40342
Sequence:	1 MATEHVNGNGTTEPPMDTSA..... RGGPSSARGVTRAGKGRGRS 545	15	880.5	30.4	586	22	AAAM38947
Scoring table:	BLOSUM62	16	789.	27.2	161	23	ABB77790
	Gapext 0.5	17	692	23.9	197	22	AAAG75683
Searched:	1107863 seqs, 158726573 residues	18	666	23.0	216	21	AAAB53301
Total number of hits satisfying chosen parameters:	1107863	19	624.5	21.5	264	20	AAV41480
Minimum DB seq length:	0	20	532.5	18.4	21	21	AAAG01393
Maximum DB seq length:	20000000000	21	411	14.2	151	22	ABB12296
Post-processing: Minimum Match 0%		22	411	14.2	151	22	AAAM40733
	Maximum Match 100%	23	395.5	13.6	182	22	AAAM53577
Listing first 45 summaries		24	395.5	13.6	182	22	AAAM42119
Database :	A_GenSeq_19Jun03:*	25	330	11.4	124	22	ABG47622
	1: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1980.DAT:*	26	330	11.4	124	22	AAW26232
	2: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1981.DAT:*	27	330	11.4	124	22	ABB32770
	3: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1982.DAT:*	28	330	11.4	124	22	ABB18253
	4: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1983.DAT:*	29	330	11.4	124	22	Human peptide enco
	5: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1984.DAT:*	30	330	11.4	124	22	Corn poly (A) bind
	6: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1985.DAT:*	31	330	11.4	124	22	Glycine max poly (
	7: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1986.DAT:*	32	330	11.4	124	22	C. albicans BAX-as
	8: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1987.DAT:*	33	330	11.4	124	22	Human secreted pro
	9: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1988.DAT:*	34	330	11.4	124	22	Rice poly (A) bind
	10: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1989.DAT:*	35	299.5	10.3	652	22	Arabidopsis thalia
	11: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1990.DAT:*	36	299.5	10.3	652	22	Arabidopsis thalia
	12: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1991.DAT:*	37	294.5	10.2	629	23	Arabidopsis thalia
	13: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1992.DAT:*	38	273	9.4	55	22	Arabidopsis thalia
	14: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1993.DAT:*	39	269	9.3	655	22	AAE13001
	15: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1994.DAT:*	40	266	9.2	503	21	AAAG46464
	16: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1995.DAT:*	41	266	9.2	515	21	AAAG46463
	17: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1996.DAT:*	42	266	9.2	547	21	AAAG46482
	18: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1997.DAT:*	43	266	9.2	559	21	AAAG46481
	19: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1998.DAT:*	44	266	9.2	618	21	AAAG46462
	20: /SIDS1/gcgdata/geneseq/geneseq/emb1/AA1999.DAT:*	45	266	9.2	662	21	AAAG46480

## ALIGNMENTS

RESULT 1

AAV24418

ID AAV24418 standard; Protein; 545 AA.

XX

AC AAV24418;

DT 21-SEP-1999 (first entry)

XX

DE Human nucleolin-like protein.

XX

KW Human; nucleolin-like protein; HNLIP; cancer; melanoma; breast cancer; prostatic cancer; autoimmune disorder; autoimmune haemolytic anaemia; inflammatory bowel disease; Alzheimer's disease.

XX

OS Homo sapiens.

XX

PN US5932475-A.

XX

PR 12-DEC-1997;

97US-0990114.

XX

PD 03-AUG-1999.

XX

PF 12-DEC-1997;

97US-0990114.

XX

PR 12-DEC-1997;

97US-0990114.

XX

PA (INCYT- INCYTE PHARM INC.

XX

PI Bandman O, Corley NC, Shah P, Yue H;

XX

DR WPI: 1999-143600/37.

XX

DN-PSDB; AA90183.

PT Nucleic acids encoding human nucleolin-like proteins useful for the

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2898	100.0	545	20 AAV24418	Human nucleolin-11
2	2898	100.0	545	23 ABG12846	Human nucleolin-11
3	2864	98.8	579	21 AAB56837	Human prostate can
4	2863	98.8	623	21 AAT70242	Human RNA-associat
5	2863	98.8	704	22 AAG74545	Human colon cancer
6	2860	98.7	561	23 ABB77791	Amino acid sequenc
7	2866	96.8	762	22 ABO4610	New human diagno
8	2355	81.3	633	21 AAY54961	Human secreted pro
24:	1647.5	56.8	397	22 AAG73837	Human colon cancer

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

PT diagnosis, prevention and treatment of disorders associated with  
 PT abnormal nucleolin-like protein expression such as cancers,  
 PT abnormal nucleolin-like protein; Alzheimer's disease  
 XX

PS Claim 8; Fig 1; 33pp; English.

CC The present sequence represents human nucleolin-like peptide, designated  
 CC HNLP, is the main protein component in the nucleus of eukaryotic  
 CC cells and is an essential part of ribosome biosynthesis and also plays  
 CC an important role in importing proteins to the nucleus. HNLP may be used  
 CC in the diagnosis, prevention and treatment of disorders associated with  
 CC abnormal expression of HNLP. For example, it may be used to treat  
 CC cancers (e.g. melanoma, breast cancer and prostate cancer), autoimmune  
 CC disorders (e.g. autoimmune haemolytic anaemia and inflammatory bowel  
 CC disease) and, in particular, Alzheimer's disease. It may also be used to  
 CC study the function of the HNLP peptides, the formation (biosynthesis) of  
 CC ribosomes and the intake of proteins into the nucleus.

XX Sequence 545 AA;

Query Match 100.0%; Score 2898; DB 20; Length 545;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-243;  
 Matches 545; Mismatches 0; Indels 0; Gaps 0;

Db 1 MATEVINGNGTEEPMDTTAVTHSENFCQLDAGLPQVAKELDEIYVAGLVAHSQDLER 60

Db 1 MATEVINGNGTEEPMDTTAVTHSENFCQLDAGLPQVAKELDEIYVAGLVAHSQDLER 60

Qy 61 AIEALKREFINEDGALAVLQQLQFDSDLSHVQNSAFQSLGQVNTYQREKQTKVADSSQGD 120

Db 61 AIEALKREFINEDGALAVLQQLQFDSDLSHVQNSAFQSLGQVNTYQREKQTKVADSSQGD 120

Qy 121 EAKTKALLERTGTYLDTVTGORKYKGPPDSVYSGQQPSVGTIELFGKIPRDLFEDELVP 180

Db 121 EAKTKALLERTGTYLDTVTGORKYKGPPDSVYSGQQPSVGTIELFGKIPRDLFEDELVP 180

Qy 181 LFEKAGPIMDLRIMMDPLTGLNLRGTYAFVFTCTKEAAQAVKLYNNEHISRGKHIGVCTSV 240

Db 181 LFEKAGPIMDLRIMMDPLTGLNLRGTYAFVFTCTKEAAQAVKLYNNEHISRGKHIGVCTSV 240

Qy 241 ANNRLFVGSPPKSKTKEQIILBEFSKTVTGLDTVLYHQDDKKNRGFPLEYEDHKTAA 300

Db 241 ANNRLFVGSPPKSKTKEQIILBEFSKTVTGLDTVLYHQDDKKNRGFPLEYEDHKTAA 300

Qy 301 QARRRLMSGKVKGWGNVGVTEWADPDEPPEMAVKVYLFVRLNANTVTEELKEAQSQ 360

Db 301 QARRRLMSGKVKGWGNVGVTEWADPDEPPEMAVKVYLFVRLNANTVTEELKEAQSQ 360

Qy 361 FGKLERVKLKDYAFHEDRDGAVKAMEMNGKDLGENETIEVFAKPPDQKKERKAAQ 420

Db 361 FGKLERVKLKDYAFHEDRDGAVKAMEMNGKDLGENETIEVFAKPPDQKKERKAAQ 420

Qy 421 QAQRNQMYDDYYGGPPHMPPTGRGRGRRGGYPPDYDYYGDHYNYRG 480

Db 421 QAQRNQMYDDYYGGPPHMPPTGRGRGRRGGYPPDYDYYGDHYNYRG 480

Qy 481 YEDPYGYEFOVGARGRGRGARGAAPSRGRAAAPPGRAGYSORGPGSARGVRAGR 540

Db 481 YEDPYGYEFOVGARGRGRGARGAAPSRGRAAAPPGRAGYSQRGGPSARGVRAGR 540

Qy 541 GRGRS 545

Db 541 GRGRS 545

Qy 646 545 AA;

Db 660 545 AA;

Qy 688 545 AA;

Db 688 545 AA;

Qy 712 545 AA;

Db 716 545 AA;

Qy 744 545 AA;

Db 744 545 AA;

Qy 776 545 AA;

Db 776 545 AA;

Qy 808 545 AA;

Db 800 545 AA;

Qy 836 545 AA;

Db 828 545 AA;

Qy 864 545 AA;

Db 856 545 AA;

Qy 892 545 AA;

Db 884 545 AA;

Qy 910 545 AA;

Db 904 545 AA;

Qy 938 545 AA;

Db 934 545 AA;

Qy 962 545 AA;

Db 958 545 AA;

Qy 980 545 AA;

Db 986 545 AA;

Qy 1008 545 AA;

Db 1020 545 AA;

Qy 1056 545 AA;

Db 1052 545 AA;

Qy 1084 545 AA;

Db 1080 545 AA;

Qy 1120 545 AA;

Db 1120 545 AA;

Qy 1156 545 AA;

Db 1152 545 AA;

Qy 1184 545 AA;

Db 1180 545 AA;

Qy 1212 545 AA;

Db 1208 545 AA;

Qy 1236 545 AA;

Db 1232 545 AA;

Qy 1264 545 AA;

Db 1260 545 AA;

Qy 1292 545 AA;

Db 1296 545 AA;

Qy 1320 545 AA;

Db 1324 545 AA;

Qy 1348 545 AA;

Db 1352 545 AA;

Qy 1376 545 AA;

Db 1372 545 AA;

Qy 1404 545 AA;

Db 1400 545 AA;

Qy 1432 545 AA;

Db 1436 545 AA;

Qy 1460 545 AA;

Db 1464 545 AA;

Qy 1492 545 AA;

Db 1488 545 AA;

Qy 1516 545 AA;

Db 1512 545 AA;

Qy 1544 545 AA;

Db 1540 545 AA;

Qy 1572 545 AA;

Db 1576 545 AA;

Qy 1600 545 AA;

Db 1592 545 AA;

Qy 1628 545 AA;

Db 1624 545 AA;

Qy 1656 545 AA;

Db 1652 545 AA;

Qy 1684 545 AA;

Db 1680 545 AA;

Qy 1712 545 AA;

Db 1708 545 AA;

Qy 1740 545 AA;

Db 1744 545 AA;

Qy 1776 545 AA;

Db 1772 545 AA;

Qy 1804 545 AA;

Db 1800 545 AA;

Qy 1832 545 AA;

Db 1836 545 AA;

Qy 1864 545 AA;

Db 1860 545 AA;

Qy 1900 545 AA;

Db 1896 545 AA;

Qy 1932 545 AA;

Db 1936 545 AA;

Qy 1964 545 AA;

Db 1960 545 AA;

Qy 2000 545 AA;

Db 1996 545 AA;

Qy 2032 545 AA;

Db 2028 545 AA;

Qy 2064 545 AA;

Db 2056 545 AA;

Qy 2100 545 AA;

Db 2084 545 AA;

Qy 2132 545 AA;

Db 2128 545 AA;

Qy 2164 545 AA;

Db 2156 545 AA;

Qy 2200 545 AA;

Db 2176 545 AA;

Qy 2232 545 AA;

Db 2228 545 AA;

Qy 2264 545 AA;

Db 2256 545 AA;

Qy 2292 545 AA;

Db 2284 545 AA;

Qy 2320 545 AA;

Db 2316 545 AA;

Qy 2352 545 AA;

Db 2340 545 AA;

Qy 2384 545 AA;

Db 2376 545 AA;

Qy 2412 545 AA;

Db 2400 545 AA;

Qy 2432 545 AA;

Db 2424 545 AA;

Qy 2464 545 AA;

Db 2456 545 AA;

Qy 2492 545 AA;

Db 2484 545 AA;

Qy 2520 545 AA;

Db 2480 545 AA;

Qy 2552 545 AA;

Db 2540 545 AA;

Qy 2584 545 AA;

Db 2540 545 AA;

Qy 2616 545 AA;

Db 2580 545 AA;

Qy 2648 545 AA;

Db 2640 545 AA;

Qy 2680 545 AA;

Db 2676 545 AA;

Qy 2712 545 AA;

Db 2704 545 AA;

Qy 2744 545 AA;

Db 2736 545 AA;

Qy 2776 545 AA;

Db 2760 545 AA;

Qy 2808 545 AA;

Db 2764 545 AA;

Qy 2840 545 AA;

Db 2836 545 AA;

Qy 2872 545 AA;

Db 2864 545 AA;

Qy 2904 545 AA;

Db 2856 545 AA;

Qy 2936 545 AA;

Db 2924 545 AA;

Qy 2968 545 AA;

Db 2956 545 AA;

Qy 3000 545 AA;

Db 2944 545 AA;

Qy 3032 545 AA;

Db 2976 545 AA;

Qy 3064 545 AA;

Db 2960 545 AA;

Qy 3096 545 AA;

Db 3036 545 AA;

Qy 3128 545 AA;

Db 3064 545 AA;

Qy 3160 545 AA;

Db 3096 545 AA;

Qy 3192 545 AA;

Db 3124 545 AA;

Qy 3224 545 AA;

Db 3156 545 AA;

Qy 3256 545 AA;

Db 3240 545 AA;

Qy 3288 545 AA;

Db 3276 545 AA;

Qy 3320 545 AA;

Db 3316 545 AA;

Qy 3352 545 AA;

Db 3344 545 AA;

Qy 3384 545 AA;

Db 3376 545 AA;

Qy 3416 545 AA;

Db 3408 545 AA;

Qy 3448 545 AA;

Db 3440 545 AA;

Qy 3480 545 AA;

Db 3472 545 AA;

Qy 3512 545 AA;

Db 3504 545 AA;

Qy 3544 545 AA;

Db 3536 545 AA;

Qy 3576 545 AA;

Db 3568 545 AA;

Qy 3608 545 AA;

Db 3596 545 AA;

Qy 3640 545 AA;

Best Local Similarity 100.0%; Pred. No. 6.2e-239; Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MATEHVNNGTTEPMOTTSAVTHSENQTLIDAGLPQKVAELDEIYVAGIVAHSDLDR 60  
 1 MATEHVNNGTTEPMOTTSAVTHSENQTLIDAGLPQKVAELDEIYVAGIVAHSDLDR 60

61 AIEALKBENEDGALAVI1QFKDSDLSHVQNKSAFLCCVMKTYTREKQGKVAQDSSKGPD 120  
 61 AIEALKBENEDGALAVI1QFKDSDLSHVQNKSAFLCCVMKTYTREKQGKVAQDSSKGPD 120

121 EAKIKALLERTGTYLDTVGOKYGGPPDSVSYGQPSVCTEIFVGKIPRDLFEDELV 180  
 121 EAKIKALLERTGTYLDTVGOKYGGPPDSVSYGQPSVCTEIFVGKIPRDLFEDELV 180

181 LPEKAGPWIWDRMLMMPLTGLNRYGAVFTCTKEAAQEAVKLYNNHEIRSGKHIGVCSV 240  
 181 LPEKAGPWIWDRMLMMPLTGLNRYGAVFTCTKEAAQEAVKLYNNHEIRSGKHIGVCSV 240

241 ANNRLFVGSIPKSSTKTEQIILFEEPSKOTEGLTDVLYHQPDDEKKNRGFCPEYEDHKTAA 300  
 241 ANNRLFVGSIPKSSTKTEQIILFEEPSKOTEGLTDVLYHQPDDEKKNRGFCPEYEDHKTAA 300

301 QARRRLMSGKVKWGNVTVWADPLDEDPEVMAKVKLFVRNLANTYTEELEKAFSQ 360  
 301 QARRRLMSGKVKWGNVTVWADPLDEDPEVMAKVKLFVRNLANTYTEELEKAFSQ 360

361 FGKLERVKLKDYAFIHFDERDGAVKAMEEENKGDLLEGENTIEIVFAKPDQKRKERAQR 420  
 361 FGKLERVKLKDYAFIHFDERDGAVKAMEEENKGDLLEGENTIEIVFAKPDQKRKERAQR 420

421 QAAKNQMDYYGGPHMPPTGRGRRGGCYPDPHYGYEDYYDGYDHYNYRG 480  
 421 QAAKNQMDYYGGPHMPPTGRGRRGGCYPDPHYGYEDYYDGYDHYNYRG 480

481 YEDPYGYEDFQVARGGRGARGAAAPSRRGAAAPPGRAGYSORGGGSAREVRAGR 540  
 481 YEDPYGYEDFQVARGGRGARGAAAPSRRGAAAPPGRAGYSORGGGSARGVRAGR 540

541 GRGRS 545  
 541 GRGRS 545

RESULT 2  
 Sequence 1, Application US/09241333  
 Patent No. 6313266  
 GENERAL INFORMATION:  
 APPLICANT: Bandman, Olga  
 APPLICANT: Yue, Henry  
 APPLICANT: Corley, Neil C.  
 APPLICANT: Shah, Purvi  
 TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/241,333  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/990,114

FILING DATE: ;  
 ATTORNEY/AGENT INFORMATION: ;  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0451 US  
 TELECOMMUNICATION INFORMATION: ;  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-845-4166  
 TELEX: ;  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 545 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE: ;  
 LIBRARY: TLYMNOT05  
 CLONE: 2809795  
 US-09-241-333-1

Query Match 100.0%; Score 2898; Db 4; Length 545;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-239;  
 Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATEHVNNGTTEPMOTTSAVTHSENQTLIDAGLPQKVAELDEIYVAGIVAHSDLDR 60  
 Db 1 MATEHVNNGTTEPMOTTSAVTHSENQTLIDAGLPQKVAELDEIYVAGIVAHSDLDR 60

Qy 61 AIEALKBENEDGALAVI1QFKDSDLSHVQNKSAFLCCVMKTYTREKQGKVAQDSSKGPD 120  
 Db 61 AIEALKBENEDGALAVI1QFKDSDLSHVQNKSAFLCCVMKTYTREKQGKVAQDSSKGPD 120

Qy 61 AIEALKBENEDGALAVI1QFKDSDLSHVQNKSAFLCCVMKTYTREKQGKVAQDSSKGPD 120

Qy 61 AIEALKBENEDGALAVI1QFKDSDLSHVQNKSAFLCCVMKTYTREKQGKVAQDSSKGPD 120

Qy 121 EAKIKALLERTGTYLDTVGOKYGGPPDSVSYGQPSVCTEIFVGKIPRDLFEDELV 180  
 Db 121 EAKIKALLERTGTYLDTVGOKYGGPPDSVSYGQPSVCTEIFVGKIPRDLFEDELV 180

Qy 181 LPEKAGPWIWDRMLMMPLTGLNRYGAVFTCTKEAAQEAVKLYNNHEIRSGKHIGVCSV 240  
 Db 181 LPEKAGPWIWDRMLMMPLTGLNRYGAVFTCTKEAAQEAVKLYNNHEIRSGKHIGVCSV 240

Qy 241 ANNRLFVGSIPKSSTKTEQIILFEEPSKOTEGLTDVLYHQPDDEKKNRGFCPEYEDHKTAA 300  
 Db 241 ANNRLFVGSIPKSSTKTEQIILFEEPSKOTEGLTDVLYHQPDDEKKNRGFCPEYEDHKTAA 300

Qy 301 QARRRLMSGKVKWGNVTVWADPLDEDPEVMAKVKLFVRNLANTYTEELEKAFSQ 360  
 Db 301 QARRRLMSGKVKWGNVTVWADPLDEDPEVMAKVKLFVRNLANTYTEELEKAFSQ 360

Qy 361 FGKLERVKLKDYAFIHFDERDGAVKAMEEENKGDLLEGENTIEIVFAKPDQKRKERAQR 420  
 Db 361 FGKLERVKLKDYAFIHFDERDGAVKAMEEENKGDLLEGENTIEIVFAKPDQKRKERAQR 420

Qy 421 QAAKNQMDYYGGPHMPPTGRGRRGGCYPDPHYGYEDYYDGYDHYNYRG 480  
 Db 421 QAAKNQMDYYGGPHMPPTGRGRRGGCYPDPHYGYEDYYDGYDHYNYRG 480

Qy 481 YEDPYGYEDFQVARGGRGARGAAAPSRRGAAAPPGRAGYSORGGGSAREVRAGR 540  
 Db 481 YEDPYGYEDFQVARGGRGARGAAAPSRRGAAAPPGRAGYSORGGGSARGVRAGR 540

Qy 541 GRGRS 545  
 Db 541 GRGRS 545

Qy 301 QARRRLMSGKVKWGNVTVWADPLDEDPEVMAKVKLFVRNLANTYTEELEKAFSQ 360  
 Db 301 QARRRLMSGKVKWGNVTVWADPLDEDPEVMAKVKLFVRNLANTYTEELEKAFSQ 360

Qy 361 FGKLERVKLKDYAFIHFDERDGAVKAMEEENKGDLLEGENTIEIVFAKPDQKRKERAQR 420  
 Db 361 FGKLERVKLKDYAFIHFDERDGAVKAMEEENKGDLLEGENTIEIVFAKPDQKRKERAQR 420

Qy 421 QAAKNQMDYYGGPHMPPTGRGRRGGCYPDPHYGYEDYYDGYDHYNYRG 480  
 Db 421 QAAKNQMDYYGGPHMPPTGRGRRGGCYPDPHYGYEDYYDGYDHYNYRG 480

Qy 481 YEDPYGYEDFQVARGGRGARGAAAPSRRGAAAPPGRAGYSORGGGSARGVRAGR 540  
 Db 481 YEDPYGYEDFQVARGGRGARGAAAPSRRGAAAPPGRAGYSORGGGSARGVRAGR 540

Qy 541 GRGRS 545  
 Db 541 GRGRS 545

RESULT 3  
 Sequence 2, Application US/09347833  
 Patent No. 6394658  
 GENERAL INFORMATION:  
 APPLICANT: Famoda, Layo O.  
 APPLICANT: Odell, Joan T.  
 TITLE OF INVENTION: Factors Involved in Gene Expression

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 09:23:45 ; Search time 21 Seconds

(without alignments) 1098.068 Million cell updates/sec

Title: US-09-978-242-1

Perfect score: 2898

Sequence: 1 MATEHVNNGNGTEEPMDTSA.....RGGPGSARGVRAGKRGGRS 545

Scoring table: BLOSUM62

Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6.ptodata/1/iaa/5A\_COMB.pep:\*

2: /cgn2\_6.ptodata/1/iaa/5B\_COMB.pep:\*

3: /cgn2\_6.ptodata/1/iaa/6A\_COMB.pep:\*

4: /cgn2\_6.ptodata/1/iaa/6B\_COMB.pep:\*

5: /cgn2\_6.ptodata/1/iaa/BetaUS\_COMB.pep:\*

6: /cgn2\_6.ptodata/1/iaa/backFiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2898	100.0	545	2	US-09-978-242-1	Sequence 1, Appli
2	2898	100.0	545	4	US-09-978-242-1	Sequence 1, Appli
3	299.5	10.3	652	3	US-09-347-833-2	Sequence 2, Appli
4	269	9.3	655	3	US-09-347-833-4	Sequence 4, Appli
5	258.5	8.9	714	2	US-09-390-114-3	Sequence 3, Appli
6	258.5	8.9	714	4	US-09-241-133-3	Sequence 3, Appli
7	223.5	7.7	336	1	US-07-667-276A-8	Sequence 8, Appli
8	209	7.2	359	1	US-07-881-015-2	Sequence 2, Appli
9	209	7.2	359	1	US-08-120-827-2	Sequence 2, Appli
10	209	7.2	359	1	US-08-478-675-2	Sequence 2, Appli
11	206.5	7.1	380	1	US-07-881-015-51	Sequence 51, Appli
12	206.5	7.1	380	1	US-08-120-827-51	Sequence 51, Appli
13	205.5	7.1	380	1	US-08-478-675-51	Sequence 5, Appli
14	205.5	7.1	675	3	US-08-173-273-5	Sequence 5, Appli
15	204.5	7.1	688	3	US-08-173-273-26	Sequence 26, Appli
16	204.5	7.1	747	3	US-08-973-273-3	Sequence 3, Appli
17	197.5	6.8	341	2	US-08-538-711A-8	Sequence 8, Appli
18	197.5	6.8	341	3	US-08-725-077-8	Sequence 8, Appli
19	197.5	6.8	341	4	US-09-542-552-8	Sequence 8, Appli
20	197.5	6.8	353	2	US-08-538-711A-7	Sequence 7, Appli
21	197.5	6.8	353	3	US-08-725-027-7	Sequence 7, Appli
22	197.5	6.8	444	1	US-09-542-552-7	Sequence 7, Appli
23	197.5	6.8	444	1	US-07-881-015-3	Sequence 3, Appli
24	197.5	6.8	444	1	US-08-120-827-3	Sequence 3, Appli
25	194.5	6.7	223	1	US-07-567-276A-10	Sequence 10, Appli
26	194.5	6.7	428	3	US-09-347-833-6	Sequence 6, Appli

RESULT 1  
US-09-990-114-1  
; Sequence 1, Application US/08990114  
; Patent No. 592475  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Yue, Henry  
; APPLICANT: Cortley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/990,114  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0451 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 545 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: TLYMNNT05  
; CLONE: 2809795  
; US-08-990-114-1  
; Query Match 100.0%; Score 2898; DB 2; Length 545;

#### ALIGNMENTS

28	184	6.3	437	2	US-08-935-450-8	Sequence 8, Appli
29	181.5	6.3	323	1	US-07-67-276A-7	Sequence 7, Appli
30	181.5	6.3	485	1	US-07-881-075-1	Sequence 1, Appli
31	181.5	6.3	485	1	US-08-120-827-1	Sequence 1, Appli
32	181.5	6.3	485	1	US-08-178-675-1	Sequence 1, Appli
33	181	6.2	759	1	US-08-676-974-1	Sequence 1, Appli
34	181	6.2	759	1	US-09-198-887-1	Sequence 1, Appli
35	181	6.2	759	1	US-09-343-443B-2	Sequence 2, Appli
36	179.5	6.2	284	1	US-08-390-858B-36	Sequence 36, Appli
37	177	6.1	656	2	US-08-9343-443B-2	Sequence 4, Appli
38	177	6.1	656	3	US-09-214-564A-4	Sequence 1, Appli
39	169.5	5.8	471	3	US-08-966-938B-1	Sequence 5, Appli
40	169.5	5.8	471	4	US-09-685-8336-1	Sequence 4, Appli
41	169	5.8	414	1	US-07-167-276A-4	Sequence 4, Appli
42	166.5	5.7	414	1	US-08-390-858B-9	Sequence 9, Appli
43	158.5	5.5	217	1	US-09-575-574-4	Sequence 4, Appli
44	147.5	5.1	162	4	US-09-252-991A-17013	Sequence 17013, A

TELEFAX: 650-845-4166  
 TELE: <Unknown>  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 545 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: TLMMNOT05  
 CLONE: 2899795  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-978-242-1

Query Match 100.0%; Score 2898; DB 9; Length 545;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-239; 0; Gaps 0;  
 Matches 545; Conservative 0; Indels 0;

Qy 1 MATHEVNGNGTEEPMDTTSAVHSNFQTLDAGLPQKVAEKLDEIYAGLVLAHSGLDER 60  
 Db 1 MATHEVNGNGTEEPMDTTSAVHSNFQTLDAGLPQKVAEKLDEIYAGLVLAHSGLDER 60

Qy 61 AIEALKEFNEDGALAVLQKFQDSLHVQNKSAFLQGVMKTYREKQTKVADSSKGPD 120  
 Db 61 AIEALKEFNEDGALAVLQKFQDSLHVQNKSAFLQGVMKTYREKQTKVADSSKGPD 120

Qy 121 EAKIKALLERGTGTLDTTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180  
 Db 121 EAKIKALLERGTGTLDTTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180

Qy 122 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180  
 Db 122 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180

Qy 181 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180  
 Db 181 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180

Qy 182 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180  
 Db 182 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180

Qy 241 ANRRLFVGSIPKSKTKQIILBEPSKTEGLTDVLYHOPDDKKRNQFCFLIEDHTAA 300  
 Db 241 ANRRLFVGSIPKSKTKQIILBEPSKTEGLTDVLYHOPDDKKRNQFCFLIEDHTAA 300

Qy 258 ANRRLFVGSIPKSKTKQIILBEPSKTEGLTDVLYHOPDDKKRNQFCFLIEDHTAA 317  
 Db 258 ANRRLFVGSIPKSKTKQIILBEPSKTEGLTDVLYHOPDDKKRNQFCFLIEDHTAA 317

Qy 301 QARRRLMSCKVKGVNGVNTVWADPIDEPPDPEVAKYKVLFTVNLANTVTEILEKAFSQ 360  
 Db 301 QARRRLMSCKVKGVNGVNTVWADPIDEPPDPEVAKYKVLFTVNLANTVTEILEKAFSQ 360

Qy 318 QARRRLMSCKVKGVNGVNTVWADPIDEPPDPEVAKYKVLFTVNLANTVTEILEKAFSQ 377  
 Db 318 QARRRLMSCKVKGVNGVNTVWADPIDEPPDPEVAKYKVLFTVNLANTVTEILEKAFSQ 377

Qy 361 FGKLERVKLKDQAFHDERGAVKAMBEENNGKDLLEGENEIEIVAKPDPDKRKRKAQR 420  
 Db 361 FGKLERVKLKDQAFHDERGAVKAMBEENNGKDLLEGENEIEIVAKPDPDKRKRKAQR 420

Qy 378 FGKLERVKLKDQAFHDERGAVKAMBEENNGKDLLEGENEIEIVAKPDPDKRKRKAQR 437  
 Db 378 FGKLERVKLKDQAFHDERGAVKAMBEENNGKDLLEGENEIEIVAKPDPDKRKRKAQR 437

Qy 421 QAAKQMDYYYYGGPAMPQPPTRGRGGRRGGYGPDDYYGTYDYYGTYDYYHNYKGG 480  
 Db 421 QAAKQMDYYYYGGPAMPQPPTRGRGGRRGGYGPDDYYGTYDYYGTYDYYHNYKGG 480

Qy 438 QAAKQMDYYYYGGPAMPQPPTRGRGGRRGGYGPDDYYGTYDYYGTYDYYHNYTRGG 497  
 Db 438 QAAKQMDYYYYGGPAMPQPPTRGRGGRRGGYGPDDYYGTYDYYGTYDYYHNYTRGG 497

Qy 481 YEDPYGYEDFQVQGRRGRRGARGAAAPSRRGAAAPPGRAGYSORGGSARGVARGK 539  
 Db 481 YEDPYGYEDFQVQGRRGRRGARGAAAPSRRGAAAPPGRAGYSORGGSARGVARGK 539

Qy 498 YEDPYGYEDFQVQGRRGRRGARGAAAPSRRGAAAPPGRAGYSORGGSARGVARGK 557  
 Db 498 YEDPYGYEDFQVQGRRGRRGARGAAAPSRRGAAAPPGRAGYSORGGSARGVARGK 557

Qy 540 ----RGRGR 544  
 Db 540 ----RGRGR 544

Qy 558 GAQQQRGRSQ 567  
 Db 558 GAQQQRGRSQ 567

RESULT 3  
 US-10-106-698-5319  
 Sequence 519, Application US/10106698

RESULT 2  
 US-09-925-300-1415  
 Sequence 1415, Application US/09925300  
 Patent No. US20020151681A1  
 GENERAL INFORMATION:  
 APPLICANT: Craig Rosen,  
 APPLICANT: Steve Ruben,  
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 FILE REFERENCE: PA101  
 CURRENT APPLICATION NUMBER: US/09/925,300  
 CURRENT FILING DATE: 2001-08-10  
 PRIOR APPLICATION NUMBER: PCT/US00/05988  
 PRIOR FILING DATE: 2000-03-08  
 PRIOR APPLICATION NUMBER: 60/124,270  
 PRIOR FILING DATE: 2000-03-11  
 NUMBER OF SEQ ID NOS: 8564  
 SOFTWARE: PatentIn Ver. 3.0  
 SEQ ID NO 5319

PRIOR FILING DATE: 1999-03-12  
 NUMBER OF SEQ ID NOS: 1890  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 1415  
 LENGTH: 579  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-925-300-1415

Query Match 98.8%; Score 2864; DB 10; Length 579;  
 Best Local Similarity 98.2%; Pred. No. 7.3e-236;  
 Matches 540; Conservative 2; Indels 6; Gaps 1;

Qy 1 MATHEVNGNGTEEPMDTTSAVHSNFQTLDAGLPQKVAEKLDEIYAGLVLAHSGLDER 60  
 Db 18 MATHEVNGNGTEEPMDTTSAVHSNFQTLDAGLPQKVAEKLDEIYAGLVLAHSGLDER 60

Qy 61 AIEALKEFNEDGALAVLQKFQDSLHVQNKSAFLQGVMKTYRQEKGTYADSSKGPD 120  
 Db 78 AIEALKEFNEDGALAVLQKFQDSLHVQNKSAFLQGVMKTYRQEKGTYADSSKGPD 120

Qy 121 EAKIKALLERGTGTLDTTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180  
 Db 138 EAKIKALLERGTGTLDTTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 180

Qy 181 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 197  
 Db 198 LFEKAGPIMDLRIMMDPLTGTGQKRYGCPPDSTSQQGQPSVGTIFVGKIPRDLFEDLYP 197

Qy 241 ANRRLFVGSIPIKSRTKQIILBEPSKTEGLTDVLYHOPDDKKRNQFCFLIEDHTAA 300  
 Db 258 ANRRLFVGSIPIKSRTKQIILBEPSKTEGLTDVLYHOPDDKKRNQFCFLIEDHTAA 300

Qy 301 QARRRLMSCKVKGVNGVNTVWADPIDEPPDPEVAKYKVLFTVNLANTVTEILEKAFSQ 360  
 Db 318 QARRRLMSCKVKGVNGVNTVWADPIDEPPDPEVAKYKVLFTVNLANTVTEILEKAFSQ 377

Qy 361 FGKLERVKLKDQAFHDERGAVKAMBEENNGKDLLEGENEIEIVAKPDPDKRKRKAQR 420  
 Db 378 FGKLERVKLKDQAFHDERGAVKAMBEENNGKDLLEGENEIEIVAKPDPDKRKRKAQR 437

Qy 421 QAAKQMDYYYYGGPAMPQPPTRGRGGRRGGYGPDDYYGTYDYYGTYDYYHNYKGG 480  
 Db 438 QAAKQMDYYYYGGPAMPQPPTRGRGGRRGGYGPDDYYGTYDYYGTYDYYHNYTRGG 497

Qy 481 YEDPYGYEDFQVQGRRGRRGARGAAAPSRRGAAAPPGRAGYSORGGSARGVARGK 539  
 Db 498 YEDPYGYEDFQVQGRRGRRGARGAAAPSRRGAAAPPGRAGYSORGGSARGVARGK 557

Qy 540 ----RGRGR 544  
 Db 558 GAQQQRGRSQ 567

RESULT 3  
 US-10-106-698-5319  
 Sequence 519, Application US/10106698

GENERAL INFORMATION:  
 APPLICANT: Ruben et al.  
 TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides  
 CURRENT APPLICATION NUMBER: PA005P1  
 CURRENT FILING DATE: 2002-03-27  
 PRIOR APPLICATION NUMBER: PCT/US00/26524  
 PRIOR FILING DATE: 2000-09-28  
 PRIOR APPLICATION NUMBER: US 60/157,137  
 PRIOR FILING DATE: 1999-09-29  
 PRIOR APPLICATION NUMBER: US 60/163,280  
 PRIOR FILING DATE: 1999-11-03  
 NUMBER OF SEQ ID NOS: 8564  
 SOFTWARE: PatentIn Ver. 3.0  
 SEQ ID NO 5319

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 09:26:11 ; Search time 38 Seconds  
(without alignments)

2965.039 Million cell updates/sec

Title: US-09-978-242-1

Perfect score: 2898

Sequence: 1 MATEHVNGNGTEEPMDITSA.....RGGPGSARGVRAKGKGRGRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6\_ptodata/2/pubpaas/us07\_pubcomb.pep;\*  
2: /cgn2\_6\_ptodata/2/pubpaas/us07\_pubcomb.pep;\*  
3: /cgn2\_6\_ptodata/2/pubpaas/us06\_pubdep;\*  
4: /cgn2\_6\_ptodata/2/pubpaas/us07\_pubcomb.pep;\*  
5: /cgn2\_6\_ptodata/2/pubpaas/us07\_pubdep;\*  
6: /cgn2\_6\_ptodata/2/pubpaas/pctus\_pubcomb.pep;\*  
7: /cgn2\_6\_ptodata/2/pubpaas/us08\_pubdep;\*  
8: /cgn2\_6\_ptodata/2/pubpaas/us08\_pubcomb.pep;\*  
9: /cgn2\_6\_ptodata/2/pubpaas/us09\_pubcomb.pep;\*  
10: /cgn2\_6\_ptodata/2/pubpaas/us09b\_pubcomb.pep;\*  
11: /cgn2\_6\_ptodata/2/pubpaas/us09c\_pubcomb.pep;\*  
12: /cgn2\_6\_ptodata/2/pubpaas/us09\_pubdep;\*  
13: /cgn2\_6\_ptodata/2/pubpaas/us10\_pubcomb.pep;\*  
14: /cgn2\_6\_ptodata/2/pubpaas/us10b\_pubcomb.pep;\*  
15: /cgn2\_6\_ptodata/2/pubpaas/us10c\_pubcomb.pep;\*  
16: /cgn2\_6\_ptodata/2/pubpaas/us10\_new\_pub.pep;\*  
17: /cgn2\_6\_ptodata/2/pubpaas/us60\_pub.pep;\*  
18: /cgn2\_6\_ptodata/2/pubpaas/us60\_pubcomb.pep;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SEQUENCES

Result No.	Score	Query Match	Length	DB ID	Description
1	2898	100.0	545	9	US-09-978-242-1
2	2864	98.8	579	10	US-09-925-300-1415
3	2863	98.8	704	15	US-10-106-698-5319
4	2860	98.7	561	10	US-09-821-687-4
5	2355	81.3	633	10	US-09-821-687-10
6	2355	81.3	633	11	US-09-374-046A-128
7	1647	56.8	397	15	US-10-106-698-4611
8	1342	46.3	673	12	US-09-949-029-100
9	1075	37.1	335	10	US-09-821-687-11
10	789	27.2	161	10	US-09-821-687-11
11	692	23.9	197	15	US-10-106-698-6457
12	666	23.0	216	9	US-09-925-299-841
13	666	23.0	216	11	US-09-929-841
14	330	11.4	124	9	US-09-864-761-33551
15	266	9.2	662	12	US-10-338-777-42

SEQUENCES

Result No.	Score	Query Match	Length	DB ID	Description
1	258	5	8.9	714	9 US-09-978-242-3
17	249.5	8.6	402	14	US-10-062-254-258
18	249.5	8.6	687	12	US-10-104-047-2651
19	244	8.4	707	12	US-10-384-569-3
20	236.5	8.2	405	14	US-10-062-254-238
21	234.5	8.1	409	14	US-10-062-254-276
22	233	8.0	314	9	US-09-849-967A-3
23	233	8.0	320	12	US-10-341-434-159
24	232	8.0	428	14	US-10-062-254-250
25	228.5	7.9	430	14	US-10-062-254-252
26	227	7.8	332	12	US-10-260A-4506
27	226.5	7.8	633	12	US-10-341-434-49
28	225	7.8	301	9	US-09-799-777-64
29	221.5	7.6	432	14	US-10-062-254-242
30	217	7.5	397	14	US-10-062-254-268
31	216	7.5	378	9	US-09-849-967A-2
32	214.5	7.4	436	14	US-10-062-254-248
33	209	7.2	359	14	US-10-025-367-23
34	208	7.2	416	14	US-10-062-254-264
35	207.5	7.2	350	14	US-10-062-254-260
36	206.5	7.1	380	14	US-10-025-367-27
37	206.5	7.1	420	14	US-10-062-254-254
38	205	7.1	117	9	US-09-764-887-194
39	205	7.1	117	15	US-10-073-961-194
40	202	7.0	494	9	US-09-833-790-234
41	201	6.9	356	14	US-10-062-254-262
42	197.5	6.8	353	11	US-09-935-642-14
43	194.5	6.7	359	14	US-10-025-367-25
44	193.5	6.7	316	14	US-10-062-254-266
45	192.5	6.6	307	14	US-10-062-254-244

ALIGNMENTS

RESULT 1

US-09-978-242-1

; Sequence 1, Application US/09978242  
; Patent No. US20020098566A1  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; Yue, Henry  
; Corley, Neil C.  
; Shah, Purvi  
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/978,242  
; FILING DATE: 15-Oct-2001  
; CLASIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 09/241,333  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/990,114  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0451 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555

Qy 474 YHNYRGGYEDPYGYED-FOVGARGRGCCRARGA-APSRGRGAAAPPRGRAGYSGORGPPG 530  
 Db 479 YHDYRGGYEDPYGYEDGAVRGSGGGRGAPPGRGAPPGRGSGORGAPPG 537

RESULT 3  
 probable RNA binding protein - *Arabidopsis thaliana*  
 N: Alternate names: protein F3C22.60  
 C: Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C: Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
 C: Accession: T49019  
 R: Purnelle, B.; Masuy, D.; Boute, M.; Mewes, H.W.; Rudd, S.; Lemcik, K.; M.  
 Submitted to the Protein Sequence Database, April 2000  
 A: Reference number: 225013  
 A: Accession: T49019  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-471 < PUR>  
 A: Cross-references: EMBL:AL353912; GSPDB:GN00061; ATSP:FAC22.60  
 A: Experimental source: cultivar Columbia; BAC clone F3C22  
 C: Genetics:  
 A: Gene: ATSP:F3C22.60  
 A: Map position: 3  
 A: Map position: 3  
 A: Introns: 120/3; 159/3; 206/3; 266/3; 336/1; 395/3; 424/2  
 Query Match 15.0%; Score 435.5; DB 2; Length 471;  
 Best Local Similarity 27.6%; Pred. No. 2, 4e-20;  
 Matches 114; Conservative 75; Mismatches 165; Indels 59; Gaps 10;  
 Qy 158 PSVGETIFVKGKIRDLFDELVPLFEXAKGPIMPLRQAFVPTCTKEAQ 217  
 Db 88 PPHGSEVYLGPIPDATEBDLKGFCGSGEVTRIMKEBDGKGAFTVFRSKDIAA 147

Query Match 17.1%; Score 495.5; DB 2; Length 521;  
 Best Local Similarity 26.2%; Pred. No. 3.9e-17;  
 Matches 149; Conservative 178; Indels 145; Gaps 17;  
 Qy 53 AHSDL-DERAIALKEFNEFDGALAVLQQPKFDSDLSHVQNSKAAPLGMVMTYQREBKQGTK 111  
 Db 17 SYSEMDDEVEEQEVEYE-----EEEDDDDDVGQNA-----BEREVD 58

Query 112 VADSSKGPDPEAKTAKAILERTSYTLDVTTGORKYKGPPP-----DSVYSG- -QQPSVGTEI 164  
 Db 59 YGDTRGGDMEDVQEEIAEDDDNHDILETADDEKPKSPIDDEREKYSHLSLPHGSEV 118

Query 165 FVGKIPRDLFDELVPLFEXAKGPID-----LRLAMDPLTCNLREYAFVTF 210  
 Db 119 FIGGLRDVGEDLRLCCEIGEFVTRTAIIFVFDLIMKDRSDSKGAYFAVF 178

Query 211 CTKEAQAEAVLYNNHEIR-----SGKHIGVCIISVANNRFLFGSIPKSKTKEQI 259  
 Db 179 KTKDVAQKATEELHSKEFKASSTANCSLSSLGSKTIRCSLSSETKRNLFIGNIPKWNTEDBF 238

Query 260 LEEFSKTVTEGTDVLYHOPDDKGNRGCFLLEYEDHKTAQAAQERRLMSGRVKWGNVGT 319  
 Db 239 RKVIEDVPGVYIELKDPTNTNRGFAFLVYNNACADYSRKMDSNFKLEGNAPT 298

Query 320 VEWADPLDPEWMAKVKYFLVRNLANTYTEELEKAFQGKLERVTK-----LKD 372  
 Db 299 VTWADPKSPTPSBSAAAQKVAKLYNNKIPENTSTEQKELQFQGKLERVTK 358

RESULT 4  
 DNZPAP  
 poladenylate-binding protein - fission yeast (Schizosaccharomyces pombe)  
 C: Species: Schizosaccharomyces pombe  
 C: Date: 14-Feb-1992 #sequence\_revision 19-Jan-2001 #text\_change 19-Jan-2001  
 C: Accession: T38950; A39720  
 R: Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 Submitted to the EMBL Data Library, May 1997  
 A: Reference number: 221819  
 A: Accession: T38950  
 A: Status: translated from GB/EMBL/DDJB  
 A: Molecule type: DNA  
 A: Residues: 1-653 < SKE>  
 A: Cross-references: EMBL:Z93396; NID:9609052; PIDN:CA08762.1; PID:92104439; GSPDB:GN00  
 A: Experimental source: strain 972h; Comid C57A7  
 R: Burd, C.G.; Matunis, E.L.; Dreyfuss, G.  
 Mol. Cell. Biol. 11, 3419-3424, 1991

Qy 373 YAFIHEDRGAVKAMEEMNGKDLEGENSEIETVAKPPDKRKERKAQRAKQNMQDDY 432  
 Db 359 FGFVHAYERSALKAVKDTEREVNGQPLEVYLAKPQAERKHDPS-----Y 405

Qy 433 YGPPHIMPPTTRGRGGRGGYGYPPDYGYEDYDGGYDQYEDQVQGARGGRGARGAAP 508  
 Db 406 SYGAAPTPAPF-----VHPTFGGFAAAYPG-----A 431

Qy 493 VGARGRGGRGARGAAPSRRGAAAP-----PRGRAGY-SQRGGP-----529  
 Db 432 MGA-GUGAGTGSQPMYGRGAMPTQMVEMLLPQPGVYLVQGMPMAAPPQRPRR 490

Qy 530 -----GSAR-----GVRAGKGRGR 542  
 Db 491 NDRNNGGSGGSGCRDNSHEDGNRGRGRY 519

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6	30	233	8.0	320	1	A44485
OM protein - protein search, using sw model		31	233	8.0	320	2	S04617
Run on:	January 28, 2004, 09:23:10 ; Search time 20 Seconds (without alignments)	32	233	8.0	707	1	DNMS
Title:	US-09-978-242-1	33	232.5	8.0	414	2	JN0866
Perfect score:	2898	34	232.5	8.0	633	1	DNXLPA
Sequence:	1 MATEHVNNGTTEPMDTSA.....RGPGSARGVAGKRGGRGRS 545	35	231	8.0	566	2	T21096
Scoring table:	BLOSUM62	36	229	7.9	694	1	DNCHNL
Scoring table:	Gapop 10.0 , Gapext 0.5	37	227.5	7.9	628	2	S44138
Searched:	283308 seqs, 96168682 residues	38	227	7.8	522	2	S52451
Total number of hits satisfying chosen parameters:	283308	39	226.5	7.8	633	1	DNHUBA
Minimum DB seq length: 0		40	225	7.8	301	2	JW0079
Maximum DB seq length: 2000000000		41	223	7.7	320	2	S30192
Post-processing: Minimum Match 0%		42	222.5	7.7	646	2	T26427
Post-processing: Maximum Match 100%		43	222	7.7	655	2	T0768
Database :	PIR 7.6:*	44	221	7.6	609	2	B84783
	1: pir1:*	45	218.5	7.5	308	2	B47369
	2: pir2:*						
	3: pir3:*						
	4: pir4:*						
RESULT 1							
T02673							
heterogeneous nuclear ribonucleoprotein R - human							
C;Species:	Homo sapiens (man)						
C;Date:	24-Mar-1999	#sequence_revision	24-Mar-1999	#text_change	08-Oct-1999		
C;Accession:	T02673						
R;Chan, E.K.L.; Mathison, D.A.; Dreyfuss, G.; Steiner, G.; Tan, E.M.; Hassett, Nucleic Acids Res. 26, 439-445, 1998							
A;Title: Molecular definition of heterogeneous nuclear ribonucleoprotein R (hnRNP R) using a reference number: Z14697; MUID: 98083170; PMID: 9421497							
A;Accession:	T02673						
A;Status: Preliminary; translated from GB/EMBL/DBJ							
A;Molecule type: mRNA							
A;Residues: 1-633 <CHA>							
A;Cross-references: EMBL:AF00364; NID:92697102; PID:92697103							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%; Score 2355; DB 2; Length 633;							
Best Local Similarity 80.2%; Pred. No. 5..1e-13;							
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;							
QY 1 MATRHVNNGNT---BEPMDTTSAVIHSENFTLIDAGLPKVAEKLDETYVAGLYVHHS 55							
Db 1 MANQ-VNGNAVQLKEEEBPM-D-TSSVTHTEHYKTLIEAGLPKVAEKLDETFQTGVLVAY 58							
Query Match 81.3%;							



Result No.	Score	Query	Match	Length	DB	ID	Description
1	2355	81.3	633	1	ROR_HUMAN	043290	homo sapien
2	287.5	9.9	653	1	PABP_SCHPO	P31209	schizosaccharomyces pombe
3	286	9.9	650	1	NUCL_XENLA	P20397	xenopus laevis
4	274	9.5	345	1	SQD_DROME	Q08473	drosophila melanogaster
5	260.5	9.0	644	1	PAB4_HUMAN	Q13310	homo sapien
6	258.5	8.9	629	1	PAB2_ARATH	P42331	arabidopsis thaliana
7	258.5	8.9	713	1	NUCL_MESAU	P08199	mesocyclospora ulei
8	248	8.6	631	1	PAB3_HUMAN	Q9h661	homo sapien
9	247	8.5	500	1	GAR2_SCHPO	P41891	schizosaccharomyces pombe
10	244	8.4	706	1	NUCL_HUMAN	P19338	homo sapien
11	242.5	8.4	636	1	PAB1_MOUSE	P23341	mus musculus
12	241.5	8.3	712	1	NUCL_RAT	P13383	rattus norvegicus
13	240.5	8.3	636	1	PAB1_HUMAN	P11940	homo sapien
14	236.5	8.2	668	1	PAB5_ARATH	Q05196	arabidopsis thaliana
15	236	8.1	576	1	PABP_YEAST	P04147	saccharomyces cerevisiae
16	234	8.1	632	1	PABP_DROME	P21187	drosophila melanogaster
17	233	8.0	319	1	NUCL_MOUSE	P49312	mus musculus
18	233	8.0	706	1	NUCL_MOUSE	P09405	mus musculus
19	232.5	8.0	414	1	NOP3_YEAST	Q01560	saccharomyces cerevisiae
20	232.5	8.0	633	1	PAB1_XENLA	P20965	xenopus laevis
21	231	8.0	319	1	ROA1_RAT	P04256	rattus norvegicus
22	229	7.9	391	1	ROA1_HUMAN	P38159	homo sapien
23	229	7.9	694	1	NUCL_CHICK	P15771	gallus gallus
24	227	7.8	522	1	PAB2_HUMAN	Q15097	homo sapien
25	227	7.7	660	1	PAB3_ARATH	Q64380	arabidopsis thaliana
26	223	7.7	319	1	ROA1_MACMU	Q28521	macaca mulatta
27	222.5	7.7	353	1	ROD_RAT	Q93j54	rattus norvegicus
28	222	7.7	371	1	ROA1_HUMAN	P03616	homo sapien
29	221.5	7.6	379	1	ROA3_MOUSE	Q8995	mus musculus
30	221	7.6	609	1	PABX_ARATH	Q93q48	arabidopsis thaliana
31	219.5	7.6	305	1	ROA0_HUMAN	Q13151	homo sapien
32	219.5	7.6	355	1	ROD_HUMAN	Q14103	homo sapien
33	218.5	7.5	424	1	S3B4_HUMAN	Q15427	homo sapien

RESULT 1						
ID	ROR_HUMAN	STANDARD;	PRT;	633 AA.	ALIGMENTS	
AC	O43390;					
DT	16-OCT-2001 (Rel. 40, Created)					
DT	16-OCT-2001 (Rel. 40, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DE	Heterogeneous nuclear ribonucleoprotein R (hnRNP R).					
GN	HNPRR.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.					
NCBI_TaxID	9606;					
RN						
RP	SEQUENCE FROM N.A. MEDLINE:98083170; PubMed=9421497;					
RA	Hassfeld W., Chan E.K.L., Mathison D.A., Portman D., Dreyfuss G., Steiner G., Tan E.M.; "Molecular definition of heterogeneous nuclear ribonucleoprotein R (hnRNP R) using autoimmune antibody: immunological relationship with hnRNP P.";					
RA	Nucleic Acids Res. 26:439-445 (1998).					
CC	-1 - FUNCTION: COMPONENT OF RIBONUCLEOSOMES, WHICH ARE COMPLEXES OF AT LEAST 20 OTHER DIFFERENT HETEROGENOUS NUCLEAR RIBONUCLEOPROTEINS (hnRNP). hnRNP PLAY AN IMPORTANT ROLE IN PROCESSING OF PRECURSOR mRNA IN THE NUCLEUS.					
CC	-1 - SUBCELLULAR LOCATION: Nuclear; nucleoplasm.					
CC	-1 - SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:licenses@isb-sib.ch">licenses@isb-sib.ch</a> ).					
CC	DR EMBL; AF000364; AAC39540.1; -.					
CC	DR HSSP; P09651; 1H11.					
CC	DR Gene3D; HGNC:5047; HNRPR.					
DR	Gene3D; HGNC:5047; HNRPR.					
DR	MIM: 607201; -.					
DR	GO:0005334; C:nucleus; TAS.					
DR	GO:0005335; F:heterogeneous nuclear ribonucleoprotein; TAS.					
DR	InterPro; IPR006535; hnRNP_R_Q.					
DR	InterPro; IPR005034; RNA_rec_mot.					
DR	Pfam; PF000716; rrm_3.					
DR	PFAM; SM03160; RRM_3.					
DR	TIGRFAMs; TIGR01648; hnRNP_R_Q; 1.					
DR	PROSITE; PS00102; RRM_3.					
DR	PROSITE; PS0030; RNP_1.					
KW	Nucleic acid binding protein; RNA binding; Ribonucleoprotein; Repeat.					
FT	DOMAIN_1 153 ASP/Glu-RICH (ACIDIC).					

Db	121 EAKIKALLERGTGYTLDLRLMMDPULTGLNRGYAFTFTCKEAAQEAVKLNNHEIRSCHKHIGCV 180	Qy	61 AIEALKKEFNEGDALAVLQOQKDSLISHVONKSAFLGCKVMTYRQEKGOTKVDASSKGPD 120
Qy	181 LFKEKAP1WDLRLMMDPULTGLNRGYAFTFTCKEAAQEAVKLNNHEIRSCHKHIGCV 240	Db	61 AIEALKKEFNEGDALAVLQOQKDSLISHVONKSAFLGCKVMTYRQEKGOTKVDASSKGPD 120
Db	181 LFKEKAP1WDLRLMMDPULTGLNRGYAFTFTCKEAAQEAVKLNNHEIRSCHKHIGCV 240	Qy	61 AIEALKKEFNEGDALAVLQOQKDSLISHVONKSAFLGCKVMTYRQEKGOTKVDASSKGPD 120
Qy	241 ANNRLFVGSI PTKSKTKQ1LBEFSKVTEGLTDVLYHQD DCKKNGRF PLEYEDHKTAA 300	Db	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Db	241 ANNRLFVGSI PTKSKTKQ1LBEFSKVTEGLTDVLYHQD DCKKNGRF PLEYEDHKTAA 300	Qy	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Qy	301 QARRRLMSGKVKWVGNVGTVEADPIEDDEPEVMAKVKVLFVRNLANTVTEILEKAFSQ 360	Db	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Db	301 QARRRLMSGKVKWVGNVGTVEADPIEDDEPEVMAKVKVLFVRNLANTVTEILEKAFSQ 360	Qy	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Qy	361 FGKLERVKKLKDQYAFIHFDERGAVKAMEEMNGKDLLEGENEIEIVFAKPPDQKERKQQR 420	Db	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Db	361 FGKLERVKKLKDQYAFIHFDERGAVKAMEEMNGKDLLEGENEIEIVFAKPPDQKERKQQR 420	Qy	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Qy	421 QAAKNQMDYYGGPMPNPPTGRGRRGGYGPDDYGYDYYGYDHYNRYGG 480	Db	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Db	421 QAAKNQMDYYGGPMPNPPTGRGRRGGYGPDDYGYDYYGYDHYNRYGG 480	Qy	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Qy	481 YEDPYGYEDFQVARGRGRRGARGAAAPSRRGAAAPPRGRAGTSQRGGPSARGVRAKG- 539	Db	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Db	481 YEDPYGYEDFQVARGRGRRGARGAAAPSRRGAAAPPRGRAGTSQRGGPSARGVRAKG- 540	Qy	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Qy	540 - - - - RGRGR 544	Db	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
Db	541 GAQQQRGRGQ 550	Qy	121 EAKIKALLERGTGYTLDVTCQRKYGPPDVSQQPSVTEIIVGKIPRDLFEBBLVP 180
<b>RESULT 2</b>			
Q9QYF4	PRELIMINARY;	PRT;	561 AA.
1D	Q9QYF4;		
AC	Q9QYF4;		
DT	01-MAY-2000 (TREMBREL 13, Created)		
DT	01-MAY-2000 (TREMBREL 13, Last sequence update)		
DT	01-MAR-2003 (TREMBREL 23, Last annotation update)		
DE	SYNCRIP protein.		
GN	NSAPI OR SYNCRIP.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OC	Mituzani A., Fukuda M., Ibarai K., Shiraiishi Y., Mikoshiba K., "SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptotagmin isoforms.", J. Biol. Chem. 275:9823-9831 (2000).		
OX	NCBI_TaxID=10090; [1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=DDY;		
RC	MEDLINE-202000483; PubMed=10734137;		
RC	PL19339; 2SXL		
RA	Mituzani A., Fukuda M., Ibarai K., Shiraiishi Y., Mikoshiba K., "SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptotagmin isoforms.", J. Biol. Chem. 275:9823-9831 (2000).		
RA	EMBL: AB0135725; BAA88342.1; -		
DR	HSSP: P19339; 2SXL		
DR	MI: 1891690; NSAPI.		
DR	InterPro: IPR006535; HnRNP_R_Q.		
DR	InterPro: IPR005044; RNA_rec_mot.		
RT	SMART: SM00360; RRM; 3.		
RT	PROSITE: PS00030; RRM_RNP_1; 2.		
RT	SEQUENCE: 561 AA; 62544_MW;		
DR	TIGRFAMS; TIGR01648; hnRNP-R_Q; 1.		
DR	PROSITE: PS00102; RRM; 3.		
DR	PROSITE: PS00030; RRM_RNP_1; 2.		
SQ	Query Match 98.7%; Score 2860; DB 11; Length 561; Best Local Similarity 98.2%; Pred. No. 1.7e-191; Mismatches 539; Conservative 2; Indels 6; Gaps 1;		
Qy	1 MATEHVNNGTTEPMDTTSAVIHSENFQTLDAQIPQKVAEKLDEIYVAGLVASLDER 60	Query Match 98.7%; Score 2860; DB 4; Length 562; Best Local Similarity 98.0%; Pred. No. 1.7e-191; Mismatches 539; Conservative 3; Indels 6; Gaps 1;	
Db	1 MATEHVNNGTTEPMDTTSAVIHSENFQTLDAQIPQKVAEKLDEIYVAGLVASLDER 60	1 MATEHVNNGTTEPMDTTSAVIHSENFQTLDAQIPQKVAEKLDEIYVAGLVASLDER 60	

Result No.	Score	Query Match	Length	DB	ID	Description	Summary
1	2861	98.7	562	11	Q8BGP1	Q8BGP1 mus musculus	Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
2	2860	98.7	561	11	Q9QYF4	Q9QYF4 mus musculus	
3	2860	98.7	562	4	Q9Y583	Q9Y583 homo sapiens	
4	2860	98.7	627	11	Q8CGC2	Q8CGC2 mus musculus	
5	2857	98.6	561	4	Q96LCL1	Q96LCL1 homo sapiens	
6	2857	98.6	623	4	060506	060506 homo sapiens	
7	2829	97.6	625	11	08B891	08B891 mus musculus	
8	2647	91.4	527	4	Q8NS99	Q8NS99 homo sapiens	
9	2640	91.1	588	4	Q96LCL2	Q96LCL2 homo sapiens	
10	2533	87.4	558	11	Q91ZRO	Q91ZRO mus musculus	
11	2487	85.8	491	11	Q8CSK6	Q8CSK6 mus musculus	
12	2343	81.6	632	11	Q8VH5	Q8VH5 mus musculus	
13	2343	81.6	636	4	Q9BV64	Q9BV64 homo sapiens	
14	2227	76.8	601	11	Q99KG1	Q99KG1 mus musculus	
15	2023	69.8	410	4	Q8IW73	Q8IW73 mus musculus	
16	1667	57.5	380	11	Q9C737	Q9C737 mus musculus	
17	1379	47.6	707	5	Q9VD18	Q9VD18 drosophila	
18	1308	45.1	529	5	Q8IN47	Q8IN47 drosophila	
19	1308	45.1	571	5	Q8IN48	Q8IN48 drosophila	
20	1300	44.9	529	5	Q95IW4	Q95IW4 drosophila	
21	1152	5.5	398	611	Q9ND1	Q9ND1 caenorhabditis elegans	
22	933	5.5	324	592	Q9XSR3	Q9XSR3 canis familiaris	
23	938	32.4	593	4	Q8NT52	Q8NT52 homo sapiens	
24	937	32.3	307	11	Q8BL132	Q8BL132 mus musculus	
25	936	32.3	590	11	Q9LWT8	Q9LWT8 mus musculus	
26	936	32.3	590	11	QBBEB8	QBBEB8 mus musculus	
27	895	30.9	586	11	Q924K3	Q924K3 rattus norvegicus	
28	893	30.8	594	11	Q923K9	Q923K9 rattus norvegicus	
29	890	30.7	524	4	Q9NRG3	Q9NRG3 homo sapiens	
30	882	30.5	586	4	Q9NZD3	Q9NZD3 homo sapiens	
31	882	30.5	594	4	Q9NG94	Q9NG94 homo sapiens	
32	880	30.4	586	4	Q9NQ93	Q9NQ93 homo sapiens	
33	877	30.3	586	4	Q9NQX9	Q9NQX9 homo sapiens	
34	876	30.2	383	11	Q8CH58	Q8CH58 rattus norvegicus	
35	876	30.2	405	11	Q8CH57	Q8CH57 rattus norvegicus	
36	859	29.7	533	4	Q8TBY0	Q8TBY0 homo sapiens	
37	829	28.6	569	4	Q9NQX8	Q9NQX8 homo sapiens	
38	513	17.7	495	10	Q9ASP6	Q9ASP6 arabidopsis thaliana	
39	495	17.1	521	10	Q23093	Q23093 arabidopsis thaliana	
40	477	16.5	472	10	Q8V2R4	Q8V2R4 oryza sativa	
41	448	15.5	431	4	Q9NXC9	Q9NXC9 homo sapiens	
42	435	15.0	471	10	Q9LXJ8	Q9LXJ8 arabidopsis thaliana	
43	386	13.3	551	5	076494	076494 dictyostelia dictyostelia	
44	376	13.0	809	10	Q8RWQ1	Q8RWQ1 arabidopsis thaliana	
45	348	12.0	353	4	Q8IXX4	Q8IXX4 homo sapiens	

source	1. .2079	/organism="unknown"	528 t
BASE COUNT	639 a	377 c	535 g
ORIGIN			
Query Match	100.0%	Score 2079; DB 6; Length 2079;	
Best Local Similarity	100.0%	Pred. No. 0; Indels 0; Gaps 0;	
Matches 2079; Conservative 0; Mismatches 0;			
Qy	1	GGGGCGCGGGCACCGGGCTGGAGCGGAACTGGATGGGGTTTGGCT 60	
Db	1	GGGGCGCGGGCACGGGCTGGAGGGAACTGGATGGGGTTTGGCT 60	
Qy	61	CCAGGGGTGAGTTGGGGATTTAACAGCTTCACTGGGGCACAGGG 120	
Db	61	CCAGGGGTGAGTTGGGGATTTAACAGCTTCACTGGGGCACAGGG 120	
Qy	121	GCAGCGGAGCAGCGGTTCCGGCAACCGGATACCATCGGACAGGTTCTCGGCC 180	
Db	121	GCAGCGGAGCAGCGGTTCCGGCAACCGGATACCATCGGACAGGTTCTCGGCC 180	
Qy	181	CAACGGGAGATCTGGAAACATGGCTACAGAACATGTTAACGGTACTGAGA 240	
Db	181	CAACGGGAGATCTGGAAACATGGCTACAGAACATGTTAACGGTACTGAGA 240	
Qy	241	GCCATGGATACTCTGGAGTTACATTGAGATTCAGACATTGGTGTATGC 300	
Db	241	GCCATGGATACTCTGGAGTTACATTGAGATTCAGACATTGGTGTATGC 300	
Qy	301	TGGTTACACAGAAAGTTCTGAAAAACTTGTATGAAATTCTGGTGTAGTTC 360	
Db	301	TGGTTACACAGAAAGTTCTGAAAAACTTGTATGAAATTCTGGTGTAGTTC 360	
Qy	361	ACATAGTGAATTAGTGAAGAGCTTGAAGCTTTAAAGAAATTCAATGAGGGTC 420	
Db	361	ACATAGTGAATTAGTGAAGAGCTTGAAGCTTTAAAGAAATTCAATGAGGGTC 420	
Qy	421	ATGGCAGTCTTCACAGTTAACAGCTTAAGACAGTCTCTCATGTTAACAAAGTC 480	
Db	421	ATGGCAGTCTTCACAGTTAACAGCTTAAGACAGTCTCTCATGTTAACAAAGTC 480	
Qy	481	CTTTTTATGTGGAGCTATGGAGACTTACAGGCAAGAAAACAGGGACCAAGTAC 540	
Db	481	CTTTTTATGTGGAGCTATGGAGACTTACAGGCAAGAAAACAGGGACCAAGTAC 540	
Qy	541	AGATTCTAGTAAAGGACCAAGTGTGGAAAGAACAGGGCAACTCTCTCTCT 600	
Db	541	AGATTCTAGTAAAGGACCAAGTGTGGAAAGAACAGGGCAACTCTCTCTCT 600	
Qy	601	CACACTGAGTGACCACTGGACCACTGAGAGTATGGAGAACCTCCGTTTA 660	
Db	601	CACACTGAGTGACCACTGGACCACTGAGAGTATGGAGAACCTCCGTTTA 660	
Qy	661	TTCACTGTAGCAGCGCTTCGTGGACTGGAGTATTGTGGAAAGATCCAAAGGATCT 720	
Db	661	TTCACTGTAGCAGCGCTTCGTGGACTGGAGTATTGTGGAAAGATCCAAAGGATCT 720	
Qy	721	ATTTGGGATGACTGCACTGACTGTCCATTATTGGAGAAAAGCTGGACCTATTCGTTCT 780	
Db	721	ATTTGGGATGACTGCACTGACTGTCCATTATTGGAGAAAAGCTGGACCTATTCGTTCT 780	
Qy	781	AATGATGGATCCACTCACTGGTCAATAGAGCTTATGGTTGTCACCTTGTACAA 840	
Db	781	AATGATGGATCCACTCACTGGTCAATAGAGCTTATGGTTGTCACCTTGTACAA 840	
Qy	841	AGAAGCAGCTAGGAGCTGTTAAACTGTATAAATCTGAAATTGTGGAAAGCTGG 900	
Db	841	AGAAGCAGCTAGGAGCTGTTAAACTGTATAAATCTGAAATTGTGGAAAGCTGG 900	
Qy	901	TATTGGTGTCTGCACTCACTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGG 960	
Db	901	TATTGGTGTCTGCACTCACTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGG 960	
Qy	961	TAAACCAGGAAACAGATTCTGAAAGAACATTAGCAAAAGTAACAGGGCTCTTACAGACGCT 1020	
Db	961	TAAACCAGGAAACAGATTCTGAAAGAACATTAGCAAAAGTAACAGGGCTCTTACAGACGCT 1020	
Qy	1021	CATTTTATACCAACCGATGACAGAAACAAACAGGGCTTTCGCTTCTGATA 1080	
Db	1021	CATTTTATACCAACCGATGACAGAAACAAACAGGGCTTTCGCTTCTGATA 1080	
Qy	1081	TGAGATCACAAAACAGTGCCTGGAGGGCTAGGGCTTATGGTAAAGTCAGCT 1140	
Db	1081	TGAGATCACAAAACAGTGCCTGGAGGGCTAGGGCTTATGGTAAAGTCAGCT 1140	
Qy	1141	CTGGGGAAATGTGGAACTGTGAATGGCTGATCTATAGAGATCTCTGATCTGAGCT 1200	
Db	1141	CTGGGGAAATGTGGAACTGTGAATGGCTGATCTATAGAGATCTCTGATCTGAGCT 1200	
Qy	1201	TATGCCAAGGTTAAAGTGTGGCTGTTAGCQAACCTTGCCAAATACTGTAACAGAAGAGAT 1260	
Db	1201	TATGCCAAGGTTAAAGTGTGGCTGTTAGCQAACCTTGCCAAATACTGTAACAGAAGAGAT 1260	
Qy	1261	TTTAGAAAGGCATTACTCAGTTGGGAAACTCTGGAAACAGGTCAGAAGGTTAAAGGATTA 1320	
Db	1261	TTTAGAAAGGCATTACTCAGTTGGGAAACTCTGGAAACAGGTCAGAAGGTTAAAGGATTA 1320	
Qy	1321	TGCTTCATTCATTGTGAGGGAGATGGTGTGCTCAAGGTATGGAGAAATGATGATGC 1380	
Db	1321	TGCTTCATTCATTGTGAGGGAGATGGTGTGCTCAAGGTATGGAGAAATGATGATGC 1380	
Qy	1381	CAAGACTTGGAGGGAGAAATTGAGATTGTTGCTGCCAAGCCACCGAGTCAAGAAAG 1440	
Db	1381	CAAGACTTGGAGGGAGAAATTGAGATTGTTGCTGCCAAGCCACCGAGTCAAGAAAG 1440	
Qy	1441	GAAGAAAGAAAAGCTCAGAGGAGAGCAGCAGCAAAATAAAATCAAATGATGAGGATTACTACTA 1500	
Db	1441	GAAGAAAGAAAAGCTCAGAGGAGAGCAGCAGCAAAATAAAATCAAATGATGAGGATTACTACTA 1500	
Qy	1501	TTATGGTCACCTCATATGCCCTCCACAAAGGGTGGGGCTGAGGGTTAGGGTGG 1560	
Db	1501	TTATGGTCACCTCATATGCCCTCCACAAAGGGTGGGGCTGAGGGTTAGGGTGG 1560	
Qy	1621	TTACCCATAACTATGTGGTGTATGGAGATCCATACTATGGTTGAGATTTCAAGT 1680	
Db	1621	TTACCCATAACTATGTGGTGTATGGAGATCCATACTATGGTTGAGATTTCAAGT 1680	
Qy	1681	TGGAGCTAGGAGGGGGGTGACAGGCAAGGGCTGGCTCATCCAGGTCTGG 1740	
Db	1681	TGGAGCTAGGAGGGGGGTGACAGGCAAGGGCTGGCTCATCCAGGTCTGG 1740	
Qy	1741	GGCTGCTCTCCCGGGTGAAGCGGGTCACTGGCTTCTGACCTGGCTCTGAGACTCA 1800	
Db	1741	GGCTGCTCTCCCGGGTGAAGCGGGTCACTGGCTTCTGACCTGGCTCTGAGACTCA 1800	
Qy	1801	AGGGTTTCCAGCAGGAAAGGGCTGAGCGGCTGGCTCTGACCTGGCTCTGAGACTCA 1860	
Db	1801	AGGGTTTCCAGCAGGAAAGGGCTGAGCGGCTGGCTCTGACCTGGCTCTGAGACTCA 1860	
Qy	1861	CTTGCTATGTGGATTACCCAGGAGCTGAGTAAAGTCAGGAAATTCAGGAA 1920	
Db	1861	CTTGCTATGTGGATTACCCAGGAGCTGAGTAAAGTCAGGAAATTCAGGAA 1920	
Qy	1921	CTTAAATAATGTCGGCTGTATAGGAGCATATTCTGAGGAAACCTTCCTATGAGACTAT 1980	
Db	1921	CTTAAATAATGTCGGCTGTATAGGAGCATATTCTGAGGAAACCTTCCTATGAGACTAT 1980	
Qy	1981	CATGGAAATAAACTGGGACATGCAACTTGGACTTGTGGATTAATCTGAAACTAAAGCTACT 2040	
Db	1981	CATGGAAATAAACTGGGACATGCAACTTGGACTTGTGGATTAATCTGAAACTAAAGCTACT 2040	
Qy	2041	AATTTTCTCTGCACTGAGCAAGTTAAACTAAAGCTACT 2079	

GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenEmh] : \* *Wattabasse* :

RESULT 1  
AR177851  
LOCUS AR177851  
DEFINITION Sequence 2 from patent US 6313266.  
ACCESSION AR177851  
VERSION AR177851.1  
VERSION G1:17922026  
COMMENT  
PAT 17-DEC-2001

SOURCE	Unknown.
ORGANISM	Unclassified
REFERENCE	1 (bases 1 to 2079)
AUTHORS	Bandim, O., Yue, H., Corley, N.C. and Shah, P.
TITLE	Human nucleolin-like protein
JOURNAL	Patent : US 613266-A 2 NOV-2001;

pred NO is the number of results predicted by chance to have a

XX	WPI: 1999-443600/37.	661	TTCAGGTCAGCACCCCTCTGTTGGACTGAGATATTGCGGGAAAGATCCCAAGAGATCT	720		
DR	P-PSDB; AAY24418.	661	TTCAGGTCAGCACCCCTCTGTTGGACTGAGATATTGCGGGAAAGATCCCAAGAGATCT	720		
XX	Nucleic acids encoding human nucleolin-like proteins useful for the diagnosis, prevention and treatment of disorders associated with abnormal nucleolin-like protein expression such as cancers, autoimmune disorders and Alzheimer's disease	721	ATTTGAGATGAACTGTTCCATTATTGAGAAAGCTGGACCTATATGGATCTTCGCTCT	780		
PT		721	ATTTGAGATGAACTGTTCCATTATTGAGAAAGCTGGACCTATATGGATCTTCGCTCT	780		
PT		781	AATGATGGATCCACTCACTGGTCATAATAGAGTTATCGGTTGACTTTTGACAAA	840		
XX	PS 1; Fig 1; 33pp; English.	781	AATGATGGATCCACTCACTGGTCATAATAGAGTTATCGGTTGACTTTTGACAAA	840		
CC	The present sequence encodes human nucleolin-like peptide, designated HNLp. HNLp is the main protein component in the nucleolus of eukaryotic cells and is an essential part of ribosome biogenesis and also plays an important role in importing proteins to the nucleus. HNLp may be used in the diagnosis, prevention and treatment of disorders associated with abnormal expression of HNLp. For example, it may be used to treat cancers (e.g. melanoma, breast cancer and prostate cancer), autoimmune disorders (e.g. autoimmune haemolytic anaemia and inflammatory bowel disease) and, in particular, Alzheimer's disease. It may also be used to study the function of the HNLp peptides, the formation (biosynthesis) of ribosomes and the intake of proteins into the nucleus.	841	AGAGAGCAGTCAGGAGGGTAAACCTGTTGAAAGACA	900		
CC		841	AGAGAGCAGTCAGGAGGGTAAACCTGTTGAAAGACA	900		
CC		901	TATTTGGTGTGATCTGATTCAGTTGGCTCTATTCTCTAACAG	960		
CC		901	TATTTGGTGTGATCTGATTCAGTTGGCTCTATTCTCTAACAG	960		
CC	Sequence 2079 BP; 639 A; 377 C; 535 G; 528 T; 0 other;	961	TAACACCAAGGAAACAGATTCTGAGAAAGTTAGCAAACTAACAGGGCTTACAGACGT	1020		
CC		961	TAACACCAAGGAAACAGATTCTGAGAAATTAGCAAACTAACAGGGCTTACAGACGT	1020		
CC	Query Match 100.0%; Score 2079; DB 20; Length 2079;	1021	CATTTTATACCAACCGATGACAAGAAAAACAGGGCTTTCCTTGAAATA	1080		
CC	Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;	1021	CATTTTATACCAACCGATGACAAGAAAAACAGGGCTTTCCTTGAAATA	1080		
CC	Matches 2079; Conservative 0; Miatches 0; Indels 0;	Qy	1081	TGAGAGTACAAAACAGTGTGCCCAGGCCAGGGCTTAGGTTAATGAGTGTAAAGTCAGGT	1140	
Db	1. GGGCGCGCGCGACCGGGAGCGCGCTCGAGGGGAGTGAACTGGATGGGTTTGCTG	60	Db	1081	TGAGAGTACAAAACAGTGTGCCCAGGCCAGGGCTTAGGTTAATGAGTGTAAAGTCAGGT	1140
Qy	61 CCAGCGCGTGAAGCTGCGCTGAGCTGGCTCACTCGCGCCGACAGCGA	120	Qy	1141	CTGGGGAAATGTTGAAACTGGTGTGAACTGGCTGATCTGATCTGAGTTCTGAGGT	1200
Db	61 CCAGCGCGTGAAGCTGCGCTGAGCTGGCTCACTCGCGCCGACAGCGA	120	Db	1141	CTGGGGAAATGTTGAAACTGGTGTGAACTGGCTGATCTGATCTGAGTTCTGAGGT	1200
Qy	121 GCAAGCGAACAGCGTTCCGGCAACCGATAACATGGACGGATTTCGCCCTCACGCC	180	Qy	1201	TATGGCAAAAGTTAAAGTGTGCTGTTGATCGCACAACCTTGCCTAGAAGATCTGTAACTGAAAGAT	1260
Db	121 GCAAGCGAACAGCGTTCCGGCAACCGATAACATGGACGGATTTCGCCCTCACGCC	180	Db	1201	TATGGCAAAAGTTAAAGTGTGCTGTTGATCGCACAACCTTGCCTAGAAGATCTGTAACTGAAAGAT	1260
Qy	181 CAACCGGGAGATCTCTGAAACATGGCTACGAAACATGTGTTAATGGTACTCTGAAAGA	240	Qy	1261	TTTAAAGGGCATTAGTCAGTTGGAAAGCTGGAAAGCTGAAAGTTAAAGGATTA	1320
Db	181 CAACCGGGAGATCTCTGAAACATGGCTACGAAACATGTGTTAATGGTACTCTGAAAGA	240	Db	1261	TTTAAAGGGCATTAGTCAGTTGGAAAGCTGGAAAGCTGAAAGCTTAAAGGATTA	1320
Qy	241 GCCCCATGGATACCTCTCTGCAAGCTTCAAGGTTACATGGCTGATGATGSC	300	Qy	1321	TGCCNTCACTCATTTGTAGGCCAATGGCTGTCAGGCTATGGAAATAATGATG	1380
Db	241 GCCCCATGGATACCTCTCTGCAAGCTTCAAGGTTACATGGCTGATGATGSC	300	Db	1321	TGCCNTCACTCATTTGTAGGCCAATGGCTGTCAGGCTATGGAAATAATGATG	1380
Qy	301 TGGTTTACACAGAAATGTGAAAAACTGTGAAATTACAGTGAATTTACATGGCAAGGTAGTTGC	360	Qy	1381	CAAAGACTTGGAGGGAGAAAATTGAAATTGTGTTTGCACGGCCACAGATGAAAG	1440
Db	301 TGGTTTACACAGAAATGTGAAAAACTGTGAAATTACAGTGAATTTACATGGCAAGGTAGTTGC	360	Db	1381	CAAAGACTTGGAGGGAGAAAATTGAAATTGTGTTTGCACGGCCACAGATGAAAG	1440
Qy	361 ACATAGTGTATTAGTGAAGAAGCTTAAAGAAATTCAATGAAGACGGTGC	420	Qy	1441	GAAGAAAGAAAAGCTCACGGCAAGCAGCAAAATAATCAAAATGTTGAGGTTGAGGTGG	1560
Db	361 ACATAGTGTATTAGTGAAGAAGCTTAAAGAAATTCAATGAAGACGGTGC	420	Db	1441	GAAGAAAGAAAAGCTCACGGCAAGCAGCAAAATAATCAAAATGTTGAGGTTGAGGTGG	1560
Qy	421 ATGGCAAGTTCTCAAGGTTAAAGCAAGCTGATCTCTGCAAGAACAGTGC	480	Qy	1501	TTATGGATCATCTCCAGATTATGGATGAAATTATGGTATGTTATGGTATGA	1620
Db	421 ATGGCAAGTTCTCAAGGTTAAAGCAAGCTGATCTCTGCAAGAACAGTGC	480	Db	1501	TTATGGATCATCTCCAGATTATGGATGAAATTATGGTATGTTATGGTATGA	1620
Qy	481 CTTTTTATGGAGTCAAGACTTACAGACTTACAGGCAAACTGGACCAAGTGC	540	Qy	1621	TTACACATAACTATCGGTTGATATGAGATCCATACTATGGTTATGAAAGATTTCAGCT	1680
Db	481 CTTTTTATGGAGTCAAGACTTACAGGCAAACTGGACCAAGTGC	540	Db	1621	TTACACATAACTATCGGTTGATATGAGATCCATACTATGGTTATGAAAGATTTCAGCT	1680
Qy	541 AGATTCAGTAAAGGGCAGATGGGCAAAATTAGGCACTCTGTCATGAAAGACGGCTA	600	Qy	1681	TGGACCTAGGAAAGGGTGGTAGAGGACCACTTCAGGTTCAAGGGTCTGTCAGGAACT	1740
Db	541 AGATTCAGTAAAGGGCAGATGGGCAAAATTAGGCACTCTGTCATGAAAGACGGCTA	600	Db	1681	TGGACCTAGGAAAGGGTGGTAGAGGACCACTTCAGGTTCAAGGGTCTGTCAGGAACT	1740
Qy	601 CACACTGTGATGTGACCACTGACAGGAAACTATGGAGGACCACTTCAGGTTA	660	Qy	1741	GGCTGCTCCCTCCCCGGTATTACAGAGGAGGGGGTATTACAGAGGAGGGTCTGTCAGGAACT	1800
Db	601 CACACTGTGATGTGACCACTGACAGGAAACTATGGAGGACCACTTCAGGTTA	660	Qy			

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

## ALIGNMENTS

Result No.	Score	Query	Match Length	DB ID	Description
1	19Jun03:*				
2:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1980.DAT:*				
3:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1981.DAT:*				
4:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1982.DAT:*				
5:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1983.DAT:*				
6:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1984.DAT:*				
7:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1985.DAT:*				
8:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1986.DAT:*				
9:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1987.DAT:*				
10:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1988.DAT:*				
11:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1989.DAT:*				
12:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1990.DAT:*				
13:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1991.DAT:*				
14:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1992.DAT:*				
15:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1993.DAT:*				
16:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1994.DAT:*				
17:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1995.DAT:*				
18:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1996.DAT:*				
19:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1997.DAT:*				
20:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1998.DAT:*				
21:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA1999.DAT:*				
22:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA2000.DAT:*				
23:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA2001.DAT:*				
24:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA2002.DAT:*				
25:	/SIDS1/gcgdata/geneseq/geneseq/geneseq -emb1/NA2003.DAT:*				

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

OS Homo sapiens.

XN PN US9932475-A.

XN PD 03-AUG-1999.

XN PR 12-DEC-1997; 97US-0990114.

XN PR 12-DEC-1997; 97US-0990114.

XN PA (INCYT) INCYTE PHARM INC.

XN PA Bandman O, Corley NC, Shah P, Yue H;

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can
3	1844	88.7	4231	21 AAF16040	Nucleotide sequenc
4	1826.2	87.8	3452	24 ABU59088	Human RNA-associat
5	1777.8	85.5	2680	21 AA251272	DNA encoding novel
6	1731	83.3	2407	23 AA568797	Human prostate tum
7	1467.4	70.4	3101	20 AA253013	Human secreted pro
8	882	42.4	2673	21 AAA16681	

Result No.	Score	Query	Match Length	DB ID	Description
1	2079	100.0	2079	20 AA990183	Human nucleolin-like Human cDNA encodin
2	2079	100.0	2079	24 AB552999	Human prostate can

Best Local Similarity 100.0*; Pred. No. 0; Mismatches 2079; Conservative 0; Indels 0; Gaps 0;		
Qy	1 GGGCGCGCGCGCACCGGCGCGAACCGGCGCTCGAGGGAGTGGAACTCGGATCGGGTTTGGT 60	Db
Db	1 GGGCGCGCGCGCACCGGCGCTCGAGGGAGTGGAACTCGGATCGGGTTTGGT 60	Qy
Qy	61 CCAGCGCGGTGAGCTTCGGGCAATTACAAAGGCGCAACTCGGCGGACAGGGA 120	Db
Db	61 CCAGCGCGGTGAGCTTCGGGCAATTACAAAGGCGCAACTCGGCGGACAGGGA 120	Qy
Qy	1221 GCAGGGAGACGGCTTCGGCAACCCGATAACATCGGAGGATTCTCGGCCCTAGCC 180	Db
Db	1221 GCAGGGAGACGGCTTCGGCAACCCGATAACATCGGAGGATTCTCGGCCCTAGCC 180	Qy
Qy	1811 CAACGGGAAATCTCGGAAACATGGCTACGAACTATGTTAATGGTACTGAA 240	Db
Db	1811 CAACGGGAAATCTCGGAAACATGGCTACGAACTATGTTAATGGTACTGAA 240	Qy
Qy	2411 GCCCCATGGATACTACTCTCGAGTTATCATTTCAGACATTGCTGATGC 300	Db
Db	2411 GCCCCATGGATACTACTCTCGAGTTATCATTTCAGACATTGCTGATGC 300	Qy
Qy	3011 TGGTTTACACAGAAAGTTGCTGAAAAACTATGATGAAATTACGGCTAGTTC 360	Db
Db	3011 TGGTTTACACAGAAAGTTGCTGAAAAACTATGATGAAACATGGTACTGAA 360	Qy
Qy	3611 ACATAGTGATTAGTAGAAAGGCTTAAGAAATTCAAGGAAGGGCTC 420	Db
Db	3611 ACATAGTGATTAGTAGAAAGGCTTAAGAAATTCAAGGAAGGGCTC 420	Qy
Qy	4211 ATGGGAGTTCTTCAACAGTTAACAGTGATCTCTCATGTTCAAGAACAAAGTC 480	Db
Db	4211 ATGGGAGTTCTTCAACAGTTAACAGTGATCTCTCATGTTCAAGAACAAAGTC 480	Qy
Qy	4811 CTTTTTATGTTGAGTCATGAGACTTACAGCCAGAGAAAAACAGGGACCAAAGTAGC 540	Db
Db	4811 CTTTTTATGTTGAGTCATGAGACTTACAGCCAGAGAAAAACAGGGACCAAAGTAGC 540	Qy
Qy	5411 AGATCTTAAGGACCAATTAGGCAACTCTGGAAGAACAGGCTA 600	Db
Db	5411 AGATCTTAAGGACCAATTAGGCAACTCTGGAAGAACAGGCTA 600	Qy
Qy	6011 CACACTGATGTGACCACTTGACAGAGGAGTATGGGACCACTCCAGATTGGTTA 660	Db
Db	6011 CACACTGATGTGACCACTTGACAGAGGAGTATGGGACCACTCCAGATTGGTTA 660	Qy
Qy	6611 TTCAAGGTCAAGCCGCTTCTGTTGGCACTGAGCATATTGTTGGAAAGATCCCAAGAGATCT 720	Db
Db	6611 TTCAAGGTCAAGCCGCTTCTGTTGGCACTGAGCATATTGTTGGAAAGATCCCAAGAGATCT 720	Qy
Qy	7211 ATTTGGGGATGAAACTGTTGCAATTATTGAGAAAGCTGGACCTATGGGATCT 780	Db
Db	7211 ATTTGGGGATGAACTGTTGCAATTATTGAGAAAGCTGGACCTATGGGATCT 780	Qy
Qy	7811 AATGATGGATCCACTCACTCACTGGTCTCAATAGGGTTATGGGTTGACAA 840	Db
Db	7811 AATGATGGATCCACTCACTCACTGGTCTCAATAGGGTTATGGGTTGACAA 840	Qy
Qy	8411 AGAAGCGAGCTCAGGGGGCTGCACTTCAGTTGGCAAAACTGGTCTGGAAAGACA 900	Db
Db	8411 AGAAGCGAGCTCAGGGGGCTGCACTTCAGTTGGCAAAACTGGTCTGGAAAGACA 900	Qy
Qy	9011 TAAACCAAGGAAAGATCTTGAAAGATTAGGCAACTTCTTCAAGAG 960	Db
Db	9011 TAAACCAAGGAAAGATCTTGAAAGATTAGGCAACTTCTTCAAGAG 960	Qy
Qy	9611 TAAACCAAGGAAAGATCTTGAAAGATTAGGCAACTTCTTCAAGAG 1020	Db
Db	9611 TAAACCAAGGAAAGATCTTGAAAGATTAGGCAACTTCTTCAAGAG 1020	Qy
Qy	1021 CATTTTATACCAACCCGATGACAAGAAAACAGAGGCTTCTGTTGAATA 1080	
1081 TGAAGATCACAAAACAGCTGCCAGGGCCTAGGTTAATGAGTGTAAAGTCAGGT 1140		
1081 TGAAGATCACAAAACAGCTGCCAGGGCCTAGGTTAATGAGTGTAAAGTCAGGT 1140		
1141 CTGGGGAAATGTTGGAACTGTTGAACTGGCTGATCTATAGAGATCTGTGATCTGTGAGGT 1200		
1141 CTGGGGAAATGTTGGAACTGTTGAACTGGCTGATCTATAGAGATCTGTGATCTGTGAGGT 1200		
1201 TATGGCAAGGTAAGAGCTGTTGACCTGGCAATTACTGTAACAGAGAGAT 1260		
1201 TATGGCAAGGTAAGAGCTGTTGACCTGGCAATTACTGTAACAGAGAGAT 1260		
1261 TTAGAAAAGGCAATTACTGCACTTGTGAGGTTATGGAGAAATGAAATGTTAAAGGATTA 1320		
1261 TTAGAAAAGGCAATTACTGCACTTGTGAGGTTATGGAGAAATGAAATGTTAAAGGATTA 1320		
1321 TCGCTTCATTCTATTGTGAGGGAGATGGTCTGAGGTCAAGGGTAAAGGTTAAAGGATTA 1380		
1321 TCGCTTCATTCTATTGTGAGGGAGATGGTCTGAGGTCAAGGGTAAAGGATTA 1380		
1381 CAAGACTTGGAGGAGAAATTGAAATTGAAATTGTTGCTCAAGCCACCGATCAGAAAG 1440		
1381 CAAGACTTGGAGGAGAAATTGAAATTGTTGCTCAAGCCACCGATCAGAAAG 1440		
1441 GAAGAAAGAAAAGCTCAGGGCAAGCAGCAAAATAAAATCAAATGTGAGGATTACTACTA 1500		
1441 GAAGAAAGAAAAGCTCAGGGCAAGCAGCAAAATAAAATCAAATGTGAGGATTACTACTA 1500		
1501 TTATGGTCCACCTCATATGCCCTCCCAACAAAGGTGAGGGCTGAGGGTCAAGGTTCAAGT 1560		
1501 TTATGGTCCACCTCATATGCCCTCCCAACAAAGGTGAGGGCTGAGGGTCAAGGTTCAAGT 1560		
1561 TTATGATACTATGTTGTTGATATGAGATTATTGTTGTTGATTTGTTGATGA 1620		
1561 TTATGATACTATGTTGTTGATATGAGATTATTGTTGATGA 1620		
1621 TTACCATAACTATGTGGTGGATATGAGATCCATACTATGGTTGAGAATTTCAGT 1680		
1621 TTACCATAACTATGTGGTGGATATGAGATCCATACTATGGTTGAGAATTTCAGT 1680		
1681 TGGAGCTAGGAGGGGGCTGAGGAGCAAGGGCTGCTGCATCCAGGGTCTGAGCTGA 1740		
1681 TGGAGCTAGGAGGGGGCTGAGGAGCAAGGGCTGCTGCATCCAGGGTCTGAGCTGA 1740		
1741 GGCTGCTCTCCCGGGTAGAGGGTTATTCAAGAGGGGGCTCTGGATCAGGCAAG 1800		
1741 GGCTGCTCTCCCGGGTAGAGGGGGCTCTGGATCAGGCAAG 1800		
1801 AGGCCTTCAGCAGGAGAAAAGGGGTGAGGGCTCTGGATCAGGCAAGGGCTGAGCTGA 1860		
1801 AGGCCTTCAGCAGGAGAAAAGGGGTGAGGGCTCTGGATCAGGCAAGGGCTGAGCTGA 1860		
1861 CTTGGTATGTGGTGGATATGAGCTTCACTGGGATCTGGGATCTGGGATCTGGG 1920		
1861 CTTGGTATGTGGTGGATATGAGCTTCACTGGGATCTGGGATCTGGG 1920		
1981 CATGGAAATCAAATAGGGACATGAACTTAACCTGGACTTTGAGTTGAGTTCTTAC 2040		
1981 CATGGAAATCAAATAGGGACATGAACTTAACCTGGACTTTGAGTTCTTAC 2040		
2041 AATTTCCTCTGCACTGGCAAGTTAAACTAAAGCTACT 2079		
2041 AATTTCCTCTGCACTGGCAAGTTAAACTAAAGCTACT 2079		

Result	No.	Score	Query	Match	Length	DB	ID	Description	Summaries
1	2079	100.0	2079	2	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
2	2079	100.0	2079	4	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
3	498.8	24.0	612	3	US-09-978-242-2	US-09-978-242-2	Sequence 654, Appli		
4	398.8	17.9	384	3	US-09-978-242-2	US-09-978-242-2	Sequence 3, Appli		
c	65.6	3.2	7218	1	US-09-978-242-2	US-09-978-242-2	Sequence 14, Appli		
6	48.4	2.3	2361	4	US-09-978-242-2	US-09-978-242-2	Sequence 28, Appli		
7	47.8	2.3	18596	3	US-09-978-242-2	US-09-978-242-2	Sequence 11, Appli		
8	47.8	2.3	18597	4	US-09-978-242-2	US-09-978-242-2	Sequence 8, Appli		
9	47.2	2.3	2277	1	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
10	47.2	2.3	2277	1	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
11	47.2	2.3	2277	1	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
12	46.4	2.2	1894	2	US-09-978-242-2	US-09-978-242-2	Sequence 7, Appli		
13	44.8	2.2	309	1	US-09-978-242-2	US-09-978-242-2	Sequence 3, Appli		
14	44.8	2.2	309	1	US-09-978-242-2	US-09-978-242-2	Sequence 3, Appli		
15	44.2	2.1	1505	1	US-09-978-242-2	US-09-978-242-2	Sequence 1, Appli		
16	44.2	2.1	4673	1	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
17	44.2	2.1	4673	1	PCT-US92-00018-1	PCT-US92-00018-1	Sequence 1, Appli		
c	43.6	2.1	10640	4	US-09-978-242-2	US-09-978-242-2	Sequence 5, Appli		
c	43.2	2.1	2042	2	US-09-978-242-2	US-09-978-242-2	Sequence 2, Appli		
c	42.6	2.0	527	3	US-09-978-242-2	US-09-978-242-2	Sequence 5, Appli		
c	42.4	2.0	5340	4	US-09-978-242-2	US-09-978-242-2	Sequence 21, Appli		
23	42.4	2.0	8920	3	US-09-978-242-2	US-09-978-242-2	Sequence 1, Appli		
24	42.2	2.0	2793	1	US-09-978-242-2	US-09-978-242-2	Sequence 1, Appli		
25	42.2	2.0	1664976	4	US-09-978-242-2	US-09-978-242-2	Sequence 1, Appli		
c	42.2	2.0	5340	4	US-09-978-242-2	US-09-978-242-2	Sequence 21, Appli		
27	41.8	2.0	5340	4	US-09-978-242-2	US-09-978-242-2	Sequence 21, Appli		

## ALIGNMENTS

Run on:	January 29, 2004, 00:15:58 ; Search time 109 Seconds (without alignments)	4818.675 Million cell updates/sec
Title:	US-09-978-242-2	
Perfect score:	2079	
Sequence:	1 GGGCGCGACGGG.....GTTATAACTAAAGCTACT 2079	
Scoring table:	IDENTITY_NUC	
Gapop:	10.0 , Gapext 1.0	
Searched:	569978 seqs, 220691566 residues	
Total number of hits satisfying chosen parameters:	1139956	
Minimum DB seq length:	0	
Maximum DB seq length:	2000000000	
Post-processing:	Minimum Match 0%	
	Maximum Match 100%	
	Listing First 45 summaries	
Database :	Issued Patents Na:*	
	1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*	
	2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*	
	3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*	
	4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*	
	5: /cgn2_6/prodata/2/ina/pctus_COMB.seq:*	
	6: /cgn2_6/prodata/2/ina/backfiles.seq:*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1	US-09-978-242-2	Sequence 2, Application US/08990114
	; Patent No. 5932475	
	; GENERAL INFORMATION:	
	; APPLICANT: Bandman, Olga	
	; APPLICANT: Yue, Henry	
	; APPLICANT: Corley, Neil C.	
	; APPLICANT: Shah, Purvi	
	; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN	
	; NUMBER OF SEQUENCES: 3	
	; CORRESPONDENCE ADDRESS:	
	; ADDRESSEE: Incyte Pharmaceuticals, Inc.	
	; STREET: 3174 Porter Drive	
	; CITY: Palo Alto	
	; STATE: CA	
	; COUNTRY: USA	
	; ZIP: 94304	
	; COMPUTER READABLE FORM:	
	; MEDIUM TYPE: Diskette	
	; COMPUTER: IBM Compatible	
	; OPERATING SYSTEM: DOS	
	; SOFTWARE: FASTSEQ FOR Windows Version 2.0	
	; CURRENT APPLICATION DATA:	
	; APPLICATION NUMBER: US/08/990,114	
	; FILING DATE: Herewith	
	; CLASSIFICATION:	
	; PRIOR APPLICATION DATA:	
	; APPLICATION NUMBER:	
	; FILING DATE:	
	; ATTORNEY/AGENT INFORMATION:	
	; NAME: Billings, Lucy J.	
	; REGISTRATION NUMBER: 36,749	
	; REFERENCE/DOCKET NUMBER: PF-0451 US	
	; TELECOMMUNICATION INFORMATION:	
	; TELEPHONE: 650-845-4166	
	; TELEX:	
	; INFORMATION FOR SEQ ID NO: 2:	
	; SEQ ID NO: 2:	
	; SOURCE/CHARACTERISTICS:	
	; LENGTH: 2079 base pairs	
	; TYPE: nucleic acid	
	; STRANDEDNESS: single	
	; TOPOLOGY: linear	
	; IMMEDIATE SOURCE:	
	; LIBRARY: TLYMNOT05	
	; CLONE: 2809795	
	; US-09-978-242-2	
	; US-09-978-242-2	
	; Query Match	100.0%
	; Score 2079;	DB 2;
	; Length 2079;	

TELEFAX: 650-845-4166  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2019 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: TLYMNOTOS  
 CLONE: 2805795  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query	Match	100.0%	Score 2079;	DB 9;	Length 2079;
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches	0;	Indels 0;
Matches 2079;	Conservative 0;	Gaps 0;			

Qy 1 GGGCGCGCGCGACCGGGGAGCTGGGCTCGAGGGAAACTGGGACTGGGCTTGGCTG 60  
 Db 1 GGGCGCGCGCGACCGGGGAGCTGGGCTCGAGGGAAACTGGGACTGGGCTTGGCTG 60  
 Qy 61 CGAGCGAGCAGCGTTCCCGCAACCCGATAACCATGGGAGGGATTCTCGGCCCTACGCC 120  
 Db 61 CGAGCGAGCAGCGTTCCCGCAACCCGATAACCATGGGAGGGATTCTCGGCCCTACGCC 120  
 Qy 121 GCAGCGAGCAGCGTTCCCGCAACCCGATAACCATGGGAGGGATTCTCGGCCCTACGCC 180  
 Db 121 GCAGCGAGCAGCGTTCCCGCAACCCGATAACCATGGGAGGGATTCTCGGCCCTACGCC 180  
 Qy 181 CAACGGGGAGACTCTGGAAACATGGCTAACGAAACATGGTAAATGGTACTGAGA 240  
 Db 181 CAACGGGGAGACTCTGGAAACATGGCTAACGAAACATGGTAAATGGTACTGAGA 240  
 Qy 241 GCCCATGGATACTACTCTGGAGCTTCGGAAACATGGCTAACGAAACATGGTACTGAGA 240  
 Db 241 GCCCATGGATACTACTCTGGAGCTTCGGAAACATGGCTAACGAAACATGGTACTGAGA 240  
 Qy 301 TGGTTTACACAGAAACTGTGTGAAAAACTGTAGTGAATAATTACGGTGCAGGGTAGTTGC 360  
 Db 301 TGGTTTACACAGAAACTGTGTGAAAAACTGTAGTGAATAATTACGGTGCAGGGTAGTTGC 360  
 Qy 361 ACATACGTGTTTACAGAAAGCTTAAAGAAATTCATGGAAAGGGTSC 420  
 Db 361 ACATACGTGTTTACAGAAAGCTTAAAGAAATTCATGGAAAGGGTSC 420  
 Qy 421 ATGGGAGTTCTCAACAGTTAACAGACTGATCTCTCATGGTCAAGAACAAAGTGC 480  
 Db 421 ATGGGAGTTCTCAACAGTTAACAGACTGATCTCTCATGGTCAAGAACAAAGTGC 480  
 Qy 481 CTTTTATGGGGATCATGGAAACTTACGGCAGAGAAAACAGGGCAGCAAAGTGC 540  
 Db 481 CTTTTATGGGGATCATGGAAACTTACGGCAGAGAAAACAGGGCAGCAAAGTGC 540  
 Qy 541 AGATTCATGAAAGGGACCAGATGGGCAAAATTAAAGGCACTCTGGAAAGAACGGCTA 600  
 Db 541 AGATTCATGAAAGGGACCAGATGGGCAAAATTAAAGGCACTCTGGAAAGAACGGCTA 600  
 Qy 601 CACACTGATGTGACCACTGGACAGAGAACTGACCTCTCATGGAGAACCCCTCAGATTCCGGTTA 660  
 Db 601 CACACTGATGTGACCACTGGACAGAGAACTGACCTCTCATGGAGAACCCCTCAGATTCCGGTTA 660  
 Qy 661 TTCAAGGTCAGAGCCCTCTGGTGGCAACTGGGACTATGGGAAAGATCCCAAGAGATCT 720  
 Db 661 TTCAAGGTCAGAGCCCTCTGGTGGCAACTGGGACTATGGGAAAGATCCCAAGAGATCT 720  
 Qy 721 ATTTGAGGATGAACTCTGTCATTGTCAGGAAAGCTGGGACTATGGGAAAGATCCCAAGAGATCT 780  
 Db 721 ATTTGAGGATGAACTCTGTCATTGTCAGGAAAGCTGGGACTATGGGAAAGATCCCAAGAGATCT 780  
 Qy 781 AATGATGGATCCACTGTCATAGGGTTATGGCTTGTCACTTTTGTACAA 840  
 Db 781 AATGATGGATCCACTGTCATAGGGTTATGGCTTGTCACTTTTGTACAA 840

Qy 961 TAAACACCAGGAAAGAGATTCTGAGAAACTGGGACTTACAGAGCTTACAGAGCT 1020  
 Db 961 TAAACACCAGGAAAGAGATTCTGAGAAACTGGGACTTACAGAGCTTACAGAGCT 1020

Qy 1021 CATTATACACCAACGGTACAGAAAAAACAGAGCTTNGCTTCTGATA 1080  
 Db 1021 CATTATACACCAACGGTACAGAAAAAACAGAGCTTNGCTTCTGATA 1080

Qy 1081 TGAGATCACAAAACAGTGCAGGGCTAGTTAATGAGTGTAAAGTCAGGT 1140  
 Db 1081 TGAGATCACAAAACAGTGCAGGGCTAGTTAATGAGTGTAAAGTCAGGT 1140

Qy 1141 CTGGGGAAATGTGGAACGTGTTGAACTGGCTGATCTGATCTGATCTGAGGT 1200  
 Db 1141 CTGGGGAAATGTGGAACGTGTTGAACTGGCTGATCTGATCTGATCTGAGGT 1200

Qy 1201 TATGCCAAAGGTAACAGTGGCTGTTGAACTGGCTGATCTGATCTGAGGT 1260  
 Db 1201 TATGCCAAAGGTAACAGTGGCTGTTGAACTGGCTGATCTGATCTGAGGT 1260

Qy 1261 TTAGAAAAGGGCATTAGTCAGTTGGAAACTCTGGAAAGGTAAAGGATTA 1320  
 Db 1261 TTAGAAAAGGCATTAGTCAGTTGGAAACTCTGGAAAGGTAAAGGATTA 1320

Qy 1321 TGCGTTCATTCTGGTGGGAGATGGTGTGTCAGGGTATGGAGAAATGATGG 1380  
 Db 1321 TGCGTTCATTCTGGTGGGAGATGGTGTGTCAGGGTATGGAGAAATGATGG 1380

Qy 1381 CAAGACTTGGGGAGAAATTGGAAATTGGTCAAGCTGGCCACCGATCAGAAAG 1440  
 Db 1381 CAAGACTTGGGGAGAAATTGGAAATTGGTCAAGCTGGCCACCGATCAGAAAG 1440

Qy 1441 GAAAGAAAGAAAAGCTCAGGGAAAGCAGCAAAATAAATGTAAGGATACTACTA 1500  
 Db 1441 GAAAGAAAGAAAAGCTCAGGGAAAGCAGCAAAATAAATGTAAGGATACTACTA 1500

Qy 1501 TTATGGTCCACCTCATGGCCCTCCACAAAGGTGAGGGTGGAGGTGG 1560  
 Db 1501 TTATGGTCCACCTCATGGCCCTCCACAAAGGTGAGGGTGGAGGTGG 1560

Qy 1561 TTATGGATATCCCTCAGATTATTATGGATTATTATGGTATGA 1620  
 Db 1561 TTATGGATATCCCTCAGATTATTATGGATTATTATGGTATGA 1620

Qy 1621 TTACCATAACTATGTGGGATCATCTATGGTTAGGATTTCACTATGGTTAGGATTTCACTA 1680  
 Db 1621 TTACCATAACTATGTGGGATCATCTATGGTTAGGATTTCACTATGGTTAGGATTTCACTA 1680

Qy 1681 TGGAGCTAGGGAGGGGGGTGAGGGCAACGGTCTGCATCAGGTTGTTG 1740  
 Db 1681 TGGAGCTAGGGAGGGGGGTGAGGGCAACGGTCTGCATCAGGTTGTTG 1740

Qy 1741 GGCTGCTCTCCCGGGTAGAGCCGGTTATTCACAGAGGGGTCTGGATCAGCAAG 1800  
 Db 1741 GGCTGCTCTCCCGGGTAGAGCCGGTTATTCACAGAGGGGTCTGGATCAGCAAG 1800

Qy 1801 AGGCCTTCAGCAGGGAAAGGGCTGAGGGCTCTGACCTGTACATGAGACTGA 1860  
 Db 1801 AGGCCTTCAGCAGGGAAAGGGCTGAGGGCTCTGACCTGTACATGAGACTGA 1860

Qy 1861 CTTGGTATCTGGGATACCCAGAGCTGGCAACTGGGAGTAATGGTAAGGAAATCAAGCA 1920  
 Db 1861 CTTGGTATCTGGGATACCCAGAGCTGGCAACTGGGAGTAATGGTAAGGAAATCAAGCA 1920

GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 01:57:54 ; Search time 532 Seconds

(without alignments)  
13976.470 Million cell updates/sec

Title: US-09-978-242-2

Perfect Score: 2079

Sequence: 1 GGGCGCGCGCACGGG.....GTTATAAAGCTACT 2079

Scoring table: IDENTITY\_NUC GapP 10.0 , Gapext 1.0

Searched: 2356869 seqs, 1788235258 residues

Total number of hits satisfying chosen parameters: 4713738

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA: \*

1: /cgn2\_6/ptodata/1/pubnra/US07\_PUBCOMB.seq;\*  
2: /cgn2\_6/ptodata/1/pubnra/US07\_PUBCOMB.seq;\*  
3: /cgn2\_6/ptodata/1/pubnra/US06\_NEW\_PUB.seq;\*  
4: /cgn2\_6/ptodata/1/pubnra/US06\_PUBCOMB.seq;\*  
5: /cgn2\_6/ptodata/1/pubnra/US07\_NEW\_PUB.seq;\*  
6: /cgn2\_6/ptodata/1/pubnra/PCUTS\_PUBCOMB.seq;\*  
7: /cgn2\_6/ptodata/1/pubnra/US08\_NEW\_PUB.seq;\*  
8: /cgn2\_6/ptodata/1/pubnra/US08\_PUBCOMB.seq;\*  
9: /cgn2\_6/ptodata/1/pubnra/US09\_PUBCOMB.seq;\*  
10: /cgn2\_6/ptodata/1/pubnra/US09\_PUBCOMB.seq;\*  
11: /cgn2\_6/ptodata/1/pubnra/US09\_PUBCOMB.seq;\*  
12: /cgn2\_6/ptodata/1/pubnra/US09\_NEW\_PUB.seq;\*  
13: /cgn2\_6/ptodata/1/pubnra/US09\_NEW\_PUB.seq;\*  
14: /cgn2\_6/ptodata/1/pubnra/US10\_PUBCOMB.seq;\*  
15: /cgn2\_6/ptodata/1/pubnra/US10\_PUBCOMB.seq;\*  
16: /cgn2\_6/ptodata/1/pubnra/US10\_NEW\_PUB.seq;\*  
17: /cgn2\_6/ptodata/1/pubnra/US60\_NEW\_PUB.seq;\*  
18: /cgn2\_6/ptodata/1/pubnra/US60\_PUBCOMB.seq;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2079	100.0	2079	9	US-09-978-242-2
2	1844	88.7	4231	10	US-09-925-300-175
3	1826.2	87.8	3452	10	US-09-821-687-3
4	1813.4	87.2	2369	15	US-10-106-698-1042
5	882	42.4	2673	11	US-09-374-046A-127
6	882	42.4	3159	14	US-10-071-766-85
c	7	871.6	41.9	2671	11
8	839.8	40.4	2073	15	US-10-198-846-13811
9	833.6	40.1	2701	15	US-10-106-698-334
10	35.1	1024	9	US-09-925-59-68	
11	730	35.1	1024	11	US-09-925-299-68
12	498.8	24.0	612	10	US-09-879-536-654
13	487	23.4	1907	9	US-09-822-849A-263
14	434.8	20.9	1176	15	US-10-198-846-9841
15	425	20.4	464	9	US-09-864-761-238

Sequence 7702, AP  
Sequence 7702, AP  
Sequence 7702, AP  
Sequence 7702, AP  
Sequence 8445, AP  
Sequence 1041, AP  
Sequence 643, App  
Sequence 2180, AP  
Sequence 301, App  
Sequence 17050, A  
Sequence 99, App1  
Sequence 310, App  
Sequence 24621, A  
Sequence 5910, AP  
Sequence 5809, AP  
Sequence 33148, A  
Sequence 414, App  
Sequence 415, App  
Sequence 291227, Sequence 291228

#### ALIGNMENTS

RESULT 1  
US-09-978-242-2  
; Sequence 2, Application US/09978242  
; Patent No. US20020098566A1  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; Yue, Henry  
; Corley, Neil C.  
; Shah, Purvi  
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FAST-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/978, 242  
; FILING DATE: 15-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 09/241, 333  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/990, 114  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0451 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555



Result	No.	Score	Query	Match length	DB	ID	Description	Summaries
5	1494	8	AK078158	71.9	Mus muscu	AK078158	Mus muscu	PRED
6	1386	6	BC021932	66.7	Homo sapi	BC021932	Homo sapi	PRED
7	1353	7	BC024283	65.1	Homo sapi	BC024283	Homo sapi	PRED
8	1343	8	BC019360	64.6	Homo sapi	BC019360	Homo sapi	PRED
9	1331	9	BC009176	64.0	Homo sapi	BC009176	Homo sapi	PRED
10	1327	10	BC015575	63.8	Homo sapi	BC015575	Homo sapi	PRED
11	983	11	AL541211	47.3	AL541211	AL541211	AL541211	PRED
12	975	12	AL541223	46.9	AL541223	AL541223	AL541223	PRED
13	921	13	AL541248	44.3	AL541248	AL541248	AL541248	PRED
14	884	14	BC026850	42.5	Mus muscu	BC026850	Mus muscu	PRED
15	883	15	BM466029	42.5	AGENCOURT	BM466029	AGENCOURT	PRED
16	866	16	BM686098	41.7	AGENCOURT	BM686098	AGENCOURT	PRED
17	862	17	BM809311	41.5	AGENCOURT	BM809311	AGENCOURT	PRED
18	844	18	BM945048	40.6	AGENCOURT	BM945048	AGENCOURT	PRED
19	838	19	BX420656	40.3	AGENCOURT	BX420656	AGENCOURT	PRED
20	836	20	BUS06813	40.2	AGENCOURT	BUS06813	AGENCOURT	PRED
21	831	21	BU194176	40.0	AGENCOURT	BU194176	AGENCOURT	PRED
22	829	22	BU183896	39.9	AGENCOURT	BU183896	AGENCOURT	PRED
23	826	23	BM451217	39.8	AGENCOURT	BM451217	AGENCOURT	PRED
24	807	24	BU845727	38.8	AGENCOURT	BU845727	AGENCOURT	PRED
25	768	25	BM543159	36.9	AGENCOURT	BM543159	AGENCOURT	PRED
c	26	757	BUS06813	36.5	AGENCOURT	BUS06813	AGENCOURT	PRED
c	27	756	BU194176	36.4	AGENCOURT	BU194176	AGENCOURT	PRED
c	28	747	BU183896	36.0	AGENCOURT	BU183896	AGENCOURT	PRED
c	29	744	BM451217	35.8	AGENCOURT	BM451217	AGENCOURT	PRED
c	30	740	BU845727	35.6	AGENCOURT	BU845727	AGENCOURT	PRED
c	31	740	BM543159	35.6	AGENCOURT	BM543159	AGENCOURT	PRED
c	32	736	BUS06813	35.4	AGENCOURT	BUS06813	AGENCOURT	PRED
c	33	729	AV298691	35.1	AGENCOURT	AV298691	AGENCOURT	PRED
c	34	728	CB182506	35.0	AGENCOURT	CB182506	AGENCOURT	PRED
c	35	721	BB731663	34.7	AGENCOURT	BB731663	AGENCOURT	PRED
c	36	719	BB731663	34.6	AGENCOURT	BB731663	AGENCOURT	PRED
c	37	715	BB402742	34.4	AGENCOURT	BB402742	AGENCOURT	PRED
c	38	712	BB0959242	34.3	AGENCOURT	BB0959242	AGENCOURT	PRED
c	39	710	BB0959242	34.2	AGENCOURT	BB0959242	AGENCOURT	PRED
c	40	709	BB0959242	34.1	AGENCOURT	BB0959242	AGENCOURT	PRED
c	41	704	BB0959242	33.9	AGENCOURT	BB0959242	AGENCOURT	PRED
c	42	700	BB0959242	33.7	AGENCOURT	BB0959242	AGENCOURT	PRED
c	43	698	BB0959242	33.7	AGENCOURT	BB0959242	AGENCOURT	PRED
c	44	698	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	45	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	46	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	47	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	48	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	49	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	50	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	51	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	52	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	53	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	54	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	55	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	56	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	57	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	58	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	59	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	60	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	61	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	62	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	63	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	64	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	65	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	66	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	67	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	68	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	69	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	70	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	71	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	72	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	73	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	74	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	75	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	76	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	77	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	78	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	79	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	80	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	81	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	82	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	83	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	84	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	85	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	86	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	87	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	88	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	89	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	90	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	91	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	92	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	93	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	94	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	95	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	96	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	97	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	98	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	99	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	100	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	101	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	102	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	103	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	104	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	105	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	106	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	107	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	108	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	109	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	110	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	111	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	112	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	113	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	114	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	115	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	116	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	117	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	118	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	119	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	120	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	121	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	122	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	123	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	124	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	125	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	126	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	127	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	128	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	129	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	130	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	131	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	132	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	133	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	134	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	135	697	BB0959242	33.6	AGENCOURT	BB0959242	AGENCOURT	PRED
c	136	697	BB0959242	33.6	AGENCOURT	BB0959242		